HANDBOOK OF

Social Psychology

BUREAU OF EDNL. & PSYL. RESEARCH
LIBRARY.

HANDBOOK OF Social Psychology

VOLUME I

THEORY AND METHOD

by

Gordon W. Allport
Bernard Berelson
Edgar F. Borgatta
Robert R. Bush
Morton Deutsch
Allen L. Edwards
Bert F. Green
Calvin S. Hall
D. O. Hebb
Roger W. Heyns
William W. Lambert
Gardner Lindzey
Ronald Lippitt
Eleanor E. Maccoby
Nathan Maccoby
Frederick Mosteller
Theodore R. Sarbin
Martin Scheerer
W. R. Thompson
John W. M. Whiting

Edited by

Gardner Lindzey
Harvard University

BUREAU OF EDNL. & PSYL. RESEARCH. : LIBRARY.

ADDISON-WESLEY PUBLISHING COMPANY, INC.
CAMBRIDGE 42, MASS.

301.1502
LIN

Copyright © *1954*
ADDISON-WESLEY PUBLISHING COMPANY, Inc.

Printed in the United States of America

ALL RIGHTS RESERVED. THIS BOOK, OR PARTS THEREOF, MAY NOT BE REPRODUCED IN ANY FORM WITHOUT WRITTEN PERMISSION OF THE PUBLISHERS.

Library of Congress Catalog No. 54-5970

Second printing — April 1956

Bureau Ednl. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated 5 · 1 · 59
Accs. No. 1268

CONTENTS

PREFACE vii

PART 1. HISTORICAL INTRODUCTION

CHAPTER 1. The Historical Background of Modern Social Psychology . . . 3
by GORDON W. ALLPORT, *Harvard University*

PART 2. CONTEMPORARY SYSTEMATIC POSITIONS

CHAPTER 2. Stimulus-Response Contiguity and Reinforcement Theory in Social Psychology 57
by WILLIAM W. LAMBERT, *Cornell University*

CHAPTER 3. Cognitive Theory 91
by MARTIN SCHEERER, *University of Kansas*

CHAPTER 4. Psychoanalytic Theory and Its Applications in the Social Sciences 143
by CALVIN S. HALL, *Western Reserve University*, and
GARDNER LINDZEY, *Harvard University*

CHAPTER 5. Field Theory in Social Psychology 181
by MORTON DEUTSCH, *New York University*

CHAPTER 6. Role Theory 223
by THEODORE R. SARBIN, *University of California*

PART 3. RESEARCH METHODS

CHAPTER 7. Experiments: Their Planning and Execution 259
by ALLEN L. EDWARDS, *University of Washington*

CHAPTER 8. Selected Quantitative Techniques 289
by FREDERICK MOSTELLER and ROBERT R. BUSH, *Harvard University*

CHAPTER 9. Attitude Measurement 335
by BERT F. GREEN, *Massachusetts Institute of Technology*

CHAPTER 10. Systematic Observational Techniques 370
by ROGER W. HEYNS and RONALD LIPPITT, *University of Michigan*

CHAPTER 11. Sociometric Measurement 405
by GARDNER LINDZEY and EDGAR F. BORGATTA, *Harvard University*

CHAPTER 12. The Interview: A Tool of Social Science 449
by ELEANOR E. MACCOBY, *Harvard University*, and
NATHAN MACCOBY, *Boston University*

CHAPTER 13. Content Analysis 488
by BERNARD BERELSON, *Ford Foundation*

CHAPTER 14. The Cross-Cultural Method 523
by JOHN W. M. WHITING, *Harvard University*

CHAPTER 15. The Social Significance of Animal Studies 532
by D. O. HEBB and W. R. THOMPSON, *McGill University*

AUTHOR INDEX 563

SUBJECT INDEX 577

PREFACE

The accelerating expansion of social psychology in the past two decades has led to an acute need for a source book more advanced than the ordinary textbook in the field but yet more focused than scattered periodical literature. Murchison's *Handbook of Social Psychology* (1935), the only previous attempt to meet this need, is out of date and out of print. It was this state of affairs that led us to assemble a book that would represent the major areas of social psychology at a level of difficulty appropriate for graduate students. In addition to serving the needs of graduate instruction, we anticipate that the volumes will be useful in advanced undergraduate courses and as a reference book for professional psychologists.

We first considered the possibility of preparing a *Handbook* three years ago. However, a final decision to proceed with the plan was not reached until the fall of 1951. During the interval we arranged an outline of topics that represented our convictions concerning the present state of social psychology. We then wrote to a large number of distinguished social psychologists asking them whether they felt our venture was likely to be professionally valuable and asking for criticisms of the outline we had prepared. The response to these letters was immensely gratifying — social psychologists as a group appear sufficiently altruistic to spend large amounts of time criticizing and commenting on a project of which they approve even though they may be unable to participate in it themselves. We also asked for specific recommendations of people who seemed best qualified to prepare the various chapters. After receiving answers we drastically revised our outline and proceeded to invite authors to prepare the various chapters. It was not until the spring of 1952 that we completed our list of contributors and even this list later underwent change. We first suggested (tongue in cheek) that the manuscripts be submitted by September 15, 1952. However, as we secretly expected, we were forced to change this due date to January 1, 1953. This "deadline" we tried hard to meet. But of course we failed and shifted our aspiration to June 15, 1953. Again we failed, although by now we were making substantial progress. By early in the fall of 1953 we had all the chapters excepting two, and the first volume was completed and in the hands of the publishers. The last two chapters were not received until early in 1954, when the second volume went to press.

Something should be said concerning the basis for the organization of the subject matter of these volumes. It became apparent early that there are many ways to subdivide social psychology but very little agreement concerning just which is best. Although we sought the advice of others, we found for almost every compelling suggestion an equally compelling countersuggestion. Thus, in the end, it was necessary to make many arbitrary decisions. So much for our knowledge that the *Handbook* could have been organized in many different ways. There is no single scheme that would satisfy all readers.

We early discovered that the subject matter was too voluminous to be contained in a single volume. Given this decision it seemed quite natural to present in one volume the chapters that dealt primarily with theoretical convictions or systematic positions, and also the methods and procedures commonly employed in social psychology. Likewise it seemed wise to present in one volume those chapters that focus upon the substantive findings and applications of social psychology.

The decision to place the historical introduction, theory, and method chapters in the first volume reflects a bias in favor of investigation that begins with an awareness of the message of the past, an attempt at theoretical relevance, and finally with a full knowledge of the procedural or measurement alternatives. All of the content of the first volume is seen, at least by the editor, as a necessary preparation for good investigation. These are the things the social psychologist should know before he lifts a single empirical finger. The second volume, then, can be seen as a justification of the contents of the first volume. Here are the empirical fruits stemming from the theories and methods summarized in the first volume.

But does this ideal scheme mirror common practice? Are the major empirical advances summarized in the second volume in reality a legitimate by-product of theoretical conceptions and sophisticated method? In fairness to science in action (as opposed to science on the books) we are afraid the answer is No. Social psychology has made its advances largely on the shoulders of random empiricists and naïve realists. Inability to distinguish between analytic and synthetic and a tendency toward reification of concepts has accompanied many of the most significant advances in this field. Who would say that those who view an attitude as a "construct" created by the investigator have made more of a contribution to this area of psychology than those who naïvely view attitudes as real and concrete entities? Thus we sorrowfully admit the organization we have imposed upon the *Handbook* may bear little relation to the path thus far trod in the development of social psychology. Nevertheless, it stands as a suggestion of the manner in which future development may well take place and as a reminder that the powerful weapon of systematic theory is now more nearly within the grasp of the wise psychologist than formerly. Where yesterday the theoretically oriented investigator and the random realist may have been on even terms, recent developments within the field may well have destroyed this equality. An approach efficient in the wilderness may be foolish in a more carefully mapped region. In summary, the precedence we give to theoretical positions reflects our conviction of the importance of theories as spurs to research, but may also represent a program for the future rather than a reflection of the past.

It must be conceded that not all areas of social psychology are covered in these volumes with equal thoroughness. Some gaps are due to the blind spots of the editor while others are the result of contributors failing to cover an area they originally agreed to cover and, in a few cases, to contributors who withdrew altogether. In spite of these shortcomings, the volumes in their present state provide the most comprehensive picture of social psychology that exists in one place today.

While deficiencies of the final product are my own responsibility, they exist in spite of a number of advisors who gave their time and energy generously throughout the venture. Of these collaborators none was nearly so important as Gordon Allport. In fairness he should be co-editor of the volume, as he contributed immeasurably both in matters of policy and in matters of detail. I owe a very special debt of gratitude to my wife Andrea for her tolerance, encouragement, and detailed assistance. Likewise of great importance is the contribution of Shirley H. Heinemann, who has been of constant help throughout the editorial process and in preparing the Index. Crucial to the success of this work were various additional colleagues who served as referees, reading individual chapters and suggesting changes and deletions. On this score I express my gratitude to Raymond Bauer, Anthony Davids, Edward E. Jones, Kaspar Naegele, David Schneider, and Walter Weiss. In addition, many of the contributors served as referees for chapters other than their own. I am indebted to E. G. Boring, S. S. Stevens, and Geraldine Stone for many helpful suggestions based on their experience in arranging the *Handbook of Experimental Psychology*. Mrs. Olga Crawford of Addison-Wesley played an indispensable role in final preparation of the manuscripts.

We wish to thank the following publishers for permission to reproduce tables, figures, and excerpts used in these volumes: Addison-Wesley Publishing Company for R. F. Bales, *Interaction Process Analysis;* Annual Reviews, Inc. for C. P. Stone and D. W. Taylor (Eds.), *Annual Review of Psychology;* Appleton-Century-Crofts for E. R. Hilgard, *Theories of Learning;* Beacon House for L. S. Cottrell and Ruth Gallagher, *Developments in Social Psychology;* Ernest Benn, Ltd., London, for G. LeBon, *The Crowd;* Cambridge University Press for W. McDougall, *The Group Mind;* Carnegie Press for H. S. Guetzkow (Ed.), *Groups, Leadership and Men;* Clark University Press for C. Murchison (Ed.), *Handbook of Social Psychology;* Columbia University Press for D. Bidney, *Theoretical Anthropology;* Doubleday & Co. for M. Cowley and B. Smith (Eds.), *Books That Changed Our Minds;* The Dryden Press and Helen Service for Marie Jahoda, M. Deutsch, and S. W. Cook, *Research Methods in Social Psychology;* The Dryden Press and Alfred McC. Lee for A. McC. Lee and N. D. Humphrey, *Race Riot;* Follett Publishing Co. for H. Werner, *Comparative Psychology of Mental Development;* Free Press for R. K. Merton, *Social Theory and Social Structure;* T. Parsons, R. F. Bales, and E. A. Shils, *Working Papers in the Theory of Action;* and Martha Wolfenstein and N. Leites, *Movies: A Psychological Study;* Grune and Stratton, Inc. for P. H. Hoch and J. Zubin (Eds.), *Psychosexual Development in Health and Disease;* Harcourt, Brace & Co. for C. Fries, *The Structure of English* and K. Koffka, *Principles of Gestalt Psychology;* Harper and Bros. for: N. W. Ackerman and Marie Jahoda, *Anti-Semitism and Emotional Disorder;* T. W. Adorno *et al., The Authoritarian Personality;* R. G. Barker and H. F. Wright, *One Boy's Day;* B. Bettelheim and M. Janowitz, *Dynamics of Prejudice;* H. Cantril and G. W. Allport, *The Psychology of Radio;* C. S. Ford and F. A. Beach, *Patterns of Sexual Behavior;* E. R. Guthrie, *The Psychology of Learning;* P. F. Lazarsfeld and F. N. Stanton (Eds.), *Communications Research, 1948-49;* J. Lee (Ed.), *The Language of Wisdom and Folly;* K. Lewin, *Resolving Social Conflicts;* K. Lewin, *Field Theory in Social Science;* R. K. Meron, *Mass Persuasion;* G. Murphy, *Personality;* J. H. Rohrer and M. Sherif, *Social Psychology at the Crossroads;* Harvard University Press for T. Parsons and E. A. Shils (Eds.), *Toward a General Theory of Action* and F. Roethlisberger and W. J. Dickson, *Management and the Worker;* Hogarth Press for S. Freud, *Group Psychology and the Analysis of the Ego* and *The Collected Papers of Sigmund Freud;* Henry Holt & Co., Inc. for G. W. Allport, *Personality;* T. Newcomb and E. L. Hartley, *Readings in Social Psychology;* R. S. Woodworth, *Experimental Psychology;* and L. L. Whyte, *The Next Development in Man;* Houghton Mifflin Co. for F. H. Allport, *Social Psychology* and C. R. Rogers, *Client-Centered Therapy;* International Universities Press for G. Róheim, *Psychoanalysis and Anthropology* and G. B. Wilbur and W. Muensterberger (Eds.), *Psychoanalysis and Culture;* Alfred A. Knopf, Inc. for C. Kluckhohn and H. A. Murray (Eds.), *Personality in Nature, Society, and Culture;* H. D. Lasswell, *Propaganda Technique in the World War;* H. D. Lasswell and Dorothy Blumenstock, *World Revolutionary Propaganda;* and W. F. White, *Rope and Faggot;* Liveright Publishing Co. for W. Köhler, *Gestalt Psychology;* McGraw-Hill Book Co. for H. G. Barnett, *Innovation: The Basis of Cultural Change;* D. Krech and R. S. Crutchfield, *Theory and Problems of Social Psychology;* K. Lewin, *Principles of Topological Psychology;* and M. L. Reymert (Ed.), *Feelings and Emotions;* The Macmillan Co. for H. Blumer and P. M. Hauser, *Movies, Delinquency, and Crime;* W. Lippmann, *Public Opinion;* R. W. Sellars, V. J. McFill, and M. Farber (Eds.), *Philosophy for the Future;* and F. K. Shuttleworth and M. A. May, *The Social Conduct and Attitudes of Movie Fans;* Methuen & Co., Ltd. for W. McDougall, *Introduction to Social Psychology* and W. McDougall, *Psychoanalysis and Social Psychology;* Noble and Noble, Inc. for W. D. Scott, *Psychology of Public Speaking;* W. W. Norton & Co. for E. H. Erikson, *Childhood and Society;* J. Piaget, *Play, Dreams, and Imitation in Childhood;* and J. Ruesch and G. Bateson, *Communication: The Social Matrix of Psychiatry;* Oliver and Boyd Ltd., Edinburgh, and Sir

Ronald A. Fisher for R. A. Fisher, *The Design of Experiments;* Oxford University Press, Inc. for H. Head, *Studies in Neurology* and H. A. Murray, *Explorations in Personality;* L. C. Page & Co. for C. Mackay, *Extraordinary Popular Delusions and the Madness of Crowds;* Philosophical Library, Inc. for P. L. Harriman (Ed.), *Encyclopedia of Psychology;* Prentice Hall, Inc. for S. E. Asch, *Social Psychology;* C. Morris, *Signs, Language, and Behavior;* and C. P. Stone (Ed.), *Comparative Psychology;* Princeton University Press for H. Cantril, *The Invasion from Mars;* C. I. Hovland *et al., Experiments on Mass Communication;* B. L. Smith, H. D. Lasswell, and R. D. Casey (Eds.), *Propaganda, Communication, and Public Opinion;* and S. A. Stouffer *et al., The American Soldier;* G. P. Putnam's Sons for Frances C. MacGregor and Margaret Mead, *Growth and Culture;* Rinehart & Co., Inc. for A. L. Edwards, *Experimental Design in Psychological Research;* A. L. Edwards, *Statistical Analysis for Students in Psychology and Education;* E. Fromm, *Escape from Freedom;* and OSS Assessment Staff, *The Assessment of Men;* The Ronald Press Co. for R. R. Blake and G. V. Ramsey (Eds.), *Perception: An Approach to Personality* and J. McV. Hunt (Ed.), *Personality and the Behavior Disorders;* Charles Scribner's Sons for C. H. Cooley, *Human Nature and the Social Order;* Simon and Schuster, Inc. for A. Einstein and L. Infeld, *The Evolution of Physics;* Society for Promoting Christian Knowledge, London, for J. C. Maxwell, *Matter and Motion;* George W. Stewart, Inc. for H. D. Lasswell and N. Leites, *Language of Politics;* University of California Press for S. M. Lipset, *Agrarian Socialism;* University of Chicago Press for A. L. Kroeber, *Anthropology Today;* S. Tax *et al., An Appraisal of Anthropology Today;* and D. Waples (Ed.), *Print, Radio and Film in a Democracy;* University of Minnesota Press for M. Deutsch and Mary E. Collins, *Interracial Housing: A Psychological Evaluation of a Social Experiment;* Watts & Co., London, for R. Firth, *Elements of Social Organization;* John Wiley & Sons, Inc. for H. Cantril, *The Psychology of Social Movements;* L. Carmichael (Ed.), *Manual of Child Psychology;* W. G. Cochran and Gertrude M. Cox, *Experimental Designs;* and S. S. Stevens (Ed.), *Handbook of Experimental Psychology;* Yale University Press for E. Cassirer, *An Essay on Man;* C. I. Hovland, I. L. Janis, and H. H. Kelley, *Communication and Persuasion;* N. E. Miller and J. Dollard, *Social Learning and Imitation;* and D. Riesman, *The Lonely Crowd.* We should also like to thank Albert W. Kennon, representative of the trustees of the McDougall estate, for permission to quote from W. McDougall, *The Energies of Men.* Thanks also go to Paul F. Lazarsfeld for permission to use excerpts from the following volumes of his authorship published by Duell, Sloan & Pearce, Inc.: *Radio Research, 1941, Radio Research, 1942-43,* and *The Peoples' Choice* by P. F. Lazarsfeld, B. Berelson, and Hazel Gaudet.

G. L.

April 1954

Part 1

Historical Introduction

CHAPTER 1

The Historical Background of Modern Social Psychology

GORDON W. ALLPORT
Harvard University

Is social psychology an ancient discipline? Or is it modern — even perhaps ultramodern? If we look only at the spectacular expansion of its researches and influence during and after World War II we are tempted to claim that social psychology is a creature of the mid-twentieth century. Like Minerva it seems to have been born full-grown. Its sudden popularity and productivity tempt us to disregard the past, to brush aside the speculations of our intellectual ancestors as so much rust clogging the gears of a new and vital science. Why bother with the "metaphysical stage" of speculation, as Comte called it, when a new era of positivism and progress has dawned?

But there is another and wiser way to view the matter. Our intellectual ancestors, for all their fumbling, were asking precisely the same questions that we are asking today. How does one generation, they wanted to know, impose its culture and its thought-forms upon the next? What happens to the mental life of the individual when he enters into association with others? And long before social psychology became a science political philosophers sought an answer to the question, *What is the social nature of man?* They well knew, as Vico (1725) observed, that "governments must conform to the nature of the men governed." Hence interest in the key problem of social psychology — man's social nature — is both ancient and persistent.

It is true that our intellectual forefathers lacked tools of precision for empirical research and that they were sometimes naïve in their theories, yet they bequeathed to us an important store of observation, hypotheses, and perspective, together with many shrewd insights that have stood the test of time. Even their errors and blind alleys are instructive. It has been well said that those who do not know history are doomed to repeat its mistakes. And the history of science shows that both the accomplishments and the blunders of one generation of scholars may become building stones in the hands of the next.

While the argument in favor of historical perspective is compelling, it is not all-compelling. It does not, for example, require us to make a conducted tour through a museum of oddities and antiquities. Nor does it hold us to a series of dust-covered portraits. A study of the history of social psychology can be justified only if it shows the relevance of historical backgrounds to present-day foregrounds. To this end the present essay lays stress not upon curios and portraits, but upon those psychological topics of current concern that receive genuinely helpful illumination from the past.

THE SOIL OF SOCIAL PSYCHOLOGY

If we ask an apparently simple question, "Who founded social psychology?" we run headfirst into one of the major problems of the science itself — the problem of social invention. Can we ever say that a single person turns the tide of history? Or that a single person ever founds a science? If we favor the "great man" type of explanation we could single out several candidates for the honor of founding social psychology: Plato, Aristotle, Hobbes, Comte, Hegel, Lazarus and Steinthal, Tarde, E. A. Ross, and others, depending on the criteria and the time-perspective we choose to adopt. A good case could be made for each of these men. Several of them have actually been called "the father of social psychology." But the truest answer to the question asserts that the roots of modern social psychology lie in the distinctive soil of western thought and civilization. Even more than most disciplines, social psychology requires a rich ground in which the tradition of free inquiry, a philosophy and ethics of democracy, and the prior rise of natural and biological science form a nourishing blend.

While the roots of social psychology lie in the intellectual soil of the whole western tradition, its present flowering is recognized to be characteristically an American phenomenon. One reason for the striking upsurge of social psychology in the U.S.A. lies in the pragmatic tradition of this country. National emergencies and conditions of social disruption provide special incentive to invent new techniques, and to strike out boldly for solutions to practical social problems. Social psychology began to flourish soon after the First World War. This event, followed by the spread of communism, by the great depression of the '30's, by the rise of Hitler, the genocide of the Jews, race riots, the Second World War and its consequent anomie, stimulated all branches of social science. A special challenge fell to social psychology. The question was asked: how is it possible to preserve the values of freedom and individual rights under conditions of mounting social strain and regimentation? Can science help provide an answer? This challenging question led to a burst of creative effort that added much to our understanding of the phenomena of leadership, public opinion, rumor, propaganda, prejudice, attitude change, morale, communication, race relations, and conflicts of value. While other nations have been confronted with the same world-wide emergency, it seems that in the United States the soil of western thought, fortified by practical meliorism, proved most fertile for the assertive growth of social psychology and related disciplines.

Practical and humanitarian motives have always played an important part in the development of social psychology, not only in America but in other lands as well. Yet there have been discordant and dissenting voices. In the opinion of Herbert Spencer in England, of Ludwig Gumplowicz in Austria, and of William Graham Sumner in the United States, it is both futile and dangerous for men to attempt to steer or to speed social change. Social evolution, they argue, requires time and obeys laws beyond the control of man. The only practical service of social science is to warn men not to interfere with the course of nature (or society). But these authors are in a minority. Most social psychologists share with Comte an optimistic view of man's chances to better his way of life. Has he not already improved his material welfare via natural science, and his health via biological science? Why should he not better his social relationships via social science? For the past century this optimistic outlook has persisted even in the face of slender accomplishment to date. Human relations seem stubbornly set. Wars have not been abolished, labor troubles have not abated, and racial tensions are not relieved. Give us time and give us money for research, the optimists say.

There is a more formal way of viewing the recent upsurge of interest in social psychology. It could be explained in terms of Comte's theory of three stages (1830, Vol. I, Ch. 1). Comte would say that only recently have the social sciences left the constraints of the first two stages, the *theological* and *metaphysical* respectively, and entered fully into the third stage of *positivism*. While Comte himself endeavored to inaugurate the third stage, it is clear that the fruit of his effort was delayed for nearly a century until the positivistic tools of experiment, statistics, survey methods, and like instruments were more adequately developed. Hornell Hart (1949) has plotted convincingly the recent upswing in the productions of social science, and argues that the recent acceleration marks the delayed entrance of social science into the era of positivism. The delay, only now ending, has resulted in a serious "cultural lag" (Ogburn, 1922) which is due to the prior entrance of the natural sciences into the era of positivism. Technological developments are dangerously far in front of social advances. Man can change matter into energy but cannot yet socially control the energy he creates. The cultural lag manifests itself in many phases of social life. Anomie is widespread.

Whether social science, under proper ethical guidance, can eventually reduce or eliminate the cultural lag may well be a question upon which human destiny depends.

DEFINING SOCIAL PSYCHOLOGY

No sharp boundaries demarcate social psychology from other social sciences. It overlaps political and economic science, cultural anthropology, and in many respects is indistinguishable from general psychology. Likewise, its tie with sociology is close. It was a sociologist, E. A. Ross, who wrote the first book to bear the title *Social Psychology* (1908). Of the dozens of textbooks subsequently published in the United States somewhat more than half have been written by psychologists, somewhat less than half by sociologists.

In spite of this apparent lack of autonomy, social psychology has its own core of theory and data and its own special viewpoint. Its focus of interest is upon the social nature of the individual person. By contrast, political science, sociology, and cultural anthropology take as their starting points the political, social, or cultural systems in which an individual person lives. It is obvious that a complete science of social relations, as Parsons and Shils (1951) point out, will embrace both the personality system and the many-sided social system.

With few exceptions, social psychologists regard their discipline as *an attempt to understand and explain how the thought, feeling, and behavior of individuals are influenced by the actual, imagined, or implied presence of other human beings.* The term "implied presence" refers to the many activities the individual carries out because of his position (role) in a complex social structure and his membership in a cultural group.

Sociology, anthropology, and political science are "higher level" disciplines; they seek inclusive laws of social structure, social change, cultural patterning. They wish to know the course of society with the individual extracted. What happens when all citizens of the United States are gradually replaced by other citizens? The English language continues to be spoken; the form of government is not fundamentally altered; the cycles of economic behavior and the existence of social classes continue to be pretty much as they were. By contrast, social psychology wishes to know how any given member of a society is affected by all the social stimuli that surround him. How does he learn his native language? Whence come the social and political attitudes that he develops? What happens when he becomes a member of a group or of a crowd?

Social psychology is above all else a branch of general psychology. Its center of emphasis is the same: human nature as localized in the person. Some writers argue that since personal mental life is always influenced by "the actual, imagined, or implied presence of others" then all psychology must be social. The point is tenable if we wish to press it. But in practice it has little value. There are many problems of human nature that need to be solved apart from social considerations: problems of psychophysics, sensory processes, emotional functions, memory span, the nature of personality integration. Social psychology overlaps general psychology, but is not identical with it.

In tracing the history of social psychology we are in the uncomfortable position of having to abstract from the total context of each author's thought those ideas that bear specifically on the social nature of the individual. We shall not, for example, examine the entire political philosophy of Bentham, or the entire sociology of Herbert Spencer. We pick out only those psychological assumptions upon which the remainder of their social theory rests. As for psychologists, we do not describe the total contribution of Wundt or McDougall or Freud to psychological science, but only those portions that are especially relevant to the explanation of social conduct.

Thus to write a history of social psychology requires a special type of light-fingeredness. One extracts a finding here, an idea there, and tries to show how these threads have woven themselves into the fabric of present-day social psychology.

SECONDARY SOURCES

The only way the student can achieve an accurate sense of history is to dig into the original writings of the leading authors of the past. But as the shadow of history lengthens it becomes increasingly difficult to consult primary sources exclusively. Secondary sources, if used with discrimination, have their uses. Not only do they save time, but they help a historian check his interpretations against those of others.

While there are many histories of sociological and political thought, there is no definitive history of social psychology. Selected portions of the subject are treated by Sprowls (1927) and by Karpf in her *American Social Psychology* (1932) and in her more recent addendum (1952). Cottrell (1941) has surveyed the decade 1930-40. None of these authors makes a sustained distinction between sociology and social psychology. The history of social psychology in Germany is reviewed by Lukasczyk (1951). Among the histories of sociology we sometimes find an antipsychological bias, as in Becker and Barnes, *Social Thought from Lore to Science* (1952). But in this book as well as in Sorokin (1928), Rice (1931), L. L. Bernard (1934), Barnes (1948), House (1936), we find helpful interpretations of the thought of past authors, including their basic psychological assumptions.

The history of cultural anthropology is still a neglected subject. Some assistance may be obtained from Haddon (1910, 1949), and even more from Kroeber and Kluckhohn (1952), who give a painstaking review of the concept of culture, allowing due weight for its psychological connotations.

Political theories have received many historical treatments, some of which include an extensive consideration of the social nature of man. A useful source in this connection is G. Engelmann, *Political Philosophy from Plato to Jeremy Bentham* (1927). This volume contains a series of pastiches that summarize each author's views of "the nature of men governed" in the author's own style. Other sources include Sabine (1937) and Catlin (1947). Catlin makes the interesting observation that while political *philosophy* has flourished for two millenia, we are now only at the beginnings of a political *science,* because the development of a science of politics would have been "too dangerous to the powers that be." The same reasoning may help to explain why it is only in recent years that the soil has been favorable for the growth of a relatively unimpeded science of social psychology; and why it is only in certain western lands that the growth has taken place.

We must not forget the history of philosophy. Until a century ago all social psychologists were at the same time philosophers, and many philosophers were social psychologists. Surveys that aid in tracing this relationship are written by Windelband (1931), Baker (1947), Russell (1945), Vaughan (1939). Historians of intellectual movements also have much to offer, e.g., M. E. Curti in *The Growth of American Thought* (1943).

To these sources must be added historical treatments of general psychology, for above all else social psychology depends upon the basic orientations and methods of experimental and theoretical psychology. In Boring (1950) there is much of value. Other historians of psychology, such as Brett (1912-1921) and Murphy (1949), devote special chapters to social psychology.

Social Psychology in Antiquity

From Plato to the time of Comte virtually all theories of the social nature of man were tied to, and incidental to, theories of the state. For this reason we may rightly say that up to a century ago social psychology was largely a branch of political philosophy. It was, of course, Plato and Aristotle who first aroused the curiosity of western man concerning his own social nature.

Plato tells us in *The Republic* that states originate because the individual is not self-sufficing. He needs the help of many. The type of government that results is due to the predominance of certain social sentiments. In one set of circumstances one of five possible forms of statecraft (and sentiments) gains the ascendancy, in different circumstances another:

Aristocracy depends upon the rule of philosophers whose dominant motives are love of wisdom and a passion for justice (the desire that each individual shall fulfill the calling for which by nature he is best fitted).

Timocracy arises when the state's guardians are motivated by vainglory, by love of power and military ambition. The resulting pattern is one of harshness toward subordinates and submission toward superiors.

Oligarchy results when through timocratic avarice a few become rich and succeed in establishing social values that regard wealth as supreme and poverty as evil.

Democracy is born in revolt against the insatiable greed of oligarchy. Such slogans as Liberty and Equality help evoke the aggressive protest of the masses against restraint.

Tyranny represents the lowest form of sovereignty and results from the excess of freedom which democracy engenders, allowing ignorant drones to gain power and to rule by caprice supported by force.

Plato felt that men formed social groups because they needed them. Thus he held a utilitarian or "social contract" view. Aristotle, on the other hand, saw the gregarious motive as something instinctive. Man is by nature a "political animal" (*Politics,* 1, ii). It is our inborn nature that leads us to positive affiliations and to a collective way of living. Yet Aristotle recognized that the social instinct varies in its force. It is not strong enough to lead to true unity within the state. The unity of a family is stronger. Stronger still is the unity (selfhood) of the individual.

Both Plato and Aristotle developed the analogy between the individual organism and society, thus giving an impetus to a doctrine that — as we shall see — has hardy supporters down to the present day.

Plato felt that the rule of the head (the philosopher-intellect) was essential to a good society. Aristotle inclined to have a higher opinion of diversified democracy. Since men differ greatly in their talents, in a democracy "one comprehends this, another that, and all, the whole." Plato would keep the peace through a rule by the elite. Aristotle would, when possible, preserve peace through an expression of the interests of all people. Plato reasoned deductively from his conception of an ideal society; Aristotle wrote inductively from the point of view of human nature as he found it. Plato was more of a rationalist; Aristotle more of an empiricist.

Because of the differences between them it has been said that every subsequent writer, philosopher, or scientist takes either a Platonic or an Aristotelian approach to the study of man and society. Perhaps the distinction is too sharp. However this may be, one thing is certain: Platonic and Aristotelian strands of thought are found in all western theory, past and present. The flavor of Comte, whom we next consider, is, as we shall see, primarily Aristotelian.

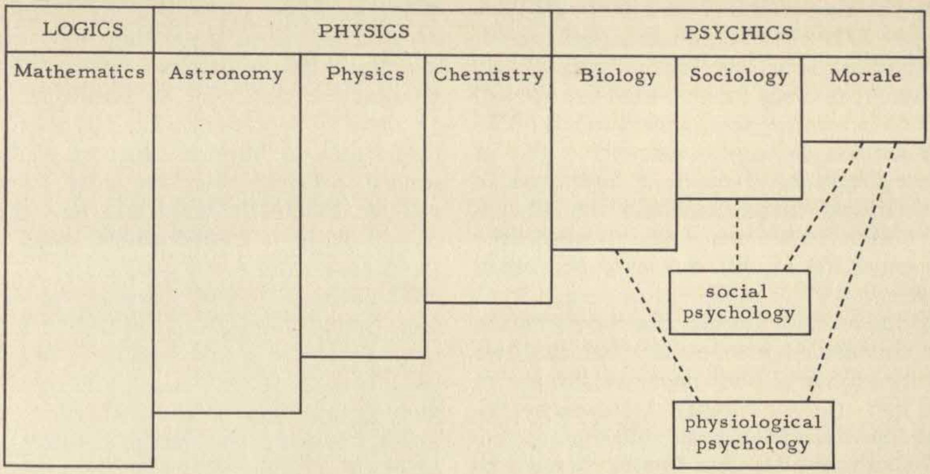


FIG. 1. Comte's hierarchy of abstract sciences. (Adapted from DeGrange, 1930, p. 211.)

FIG. 1. Comte's hierarchy of abstract sciences. (Adapted from DeGrange, 1930, p. 211.)

COMTE'S DISCOVERY OF SOCIAL PSYCHOLOGY

Two facts concerning the contributions of Auguste Comte, 1798-1857, are known to most students. First, that in his famous "law of three stages" he called attention to the gradual emergence of the sciences from the theological, through the metaphysical, into the positive stages of development. We have already commented on the bearing of this law upon the current problem of the cultural lag. Secondly, most students know that Comte classified the basic abstract sciences, and in 1839 reached the conclusion that sociology (the name is his invention) should be recognized as a new positive science. One might say that Comte christened sociology many years before it was born, but nonetheless clearly foresaw its development.

What most students do not know is that toward the end of his life Comte was wrestling with a "true final science" which, if fully worked out, could only have been what today we call psychology, even though Comte preferred to label it the science of "la morale positive." In Volume 4 of his *Système de Politique Positive* (1854) he states his intention to publish in 1859 *Le Système de morale positive, ou Traité de l'éducation universelle*. But two years before this projected date he died, leaving only fragmentary expositions of the new science. These fragments have been collated and interpreted by De Grange (1923, 1930).

What is important for us today is the logic of Comte's argument. No writer has ever given such sustained attention as he to the ordering of the sciences. The fact that he found himself finally driven to postulate a "true final science" — and that his conception of that science parallels our present conception of modern psychology (especially social psychology) — is a neglected event of historical importance. If it were possible to designate a single deliberate "founder" of social psychology as a science we should have to nominate Comte for this honor.

The reason why Comte refused to adopt the term "psychology" for his science is itself revealing (see Lévy-Bruhl, 1903, pp. 188-209). In his day, "psychology" was too rationalistic, too introspective, and too mentalistic — in short, too "metaphysical" for his taste. He feared that by preserving the name he would retard the development of the positive science that he was seeking. Just as he coined "sociology" to aid him in defining a new scientific field, so too he coined "la morale" to designate the final science which should follow and be built upon biology and sociology.

The abstract sciences. Comte insists upon separating the abstract sciences from the concrete. The former deal with irreducible phenomena, with elements and their analysis, with "events" that are fundamental and primary; the latter deal with composite phenomena, with concrete "beings" (not abstract events), and with applications of the abstract sciences. Figure 1 lists the seven abstract sciences, in the order of their presumed emergence. All other sciences are concrete; among them are geology, meteorology, botany, zoology, and education (the last being an example of applied morale).

The nature of these sciences is described in detail in two long series of volumes, the *Positive Philosophy,* 1830-1842, and *Positive Polity,* 1851-1854. The third series, the *Subjective Synthesis,* which was to include in Volume Two his full account of la morale, was never finished. The abstract sciences, he argued, can evolve only in a definite hierarchical order, as indicated in Fig. 1. Mankind can progress only from what is known to what is unknown, and only now is the vista opening for a positive science of human behavior.

It was the level of Psychics that gave him the greatest trouble. At first he discerned only two irreducible classes of phenomena pertaining to sentient life: (1) the biological, organic, or vital, and (2) the sociological, collective, or societal. Gradually he felt himself forced to recognize a third class of phenomena: (3) the moral or individual, wherein the affective, intellectual, and active functions of the person are bound together.

The nature of la morale. As early as the second volume of the *Positive Polity* (1852) Comte asked himself the momentous question: "How can the individual be at once cause and consequence of society?" A product of society, he is yet a creative agent, not to be understood in terms of the gross generalizations of sociology. While an individual is a concrete being, he requires an abstract science that will examine his nature from a point of view that grows out of, but is yet beyond the biological and the sociological points of view. Man is something more than a biological being, something more than a "cultural cumulation" with which sociology deals. He is a moral agent, uniquely patterned and deserving a special science to explore and interpret his nature.

> When biology is looked upon as merely outlining the study of human existence in the light of the study of the vegetative and animal functions, then sociology alone gives a knowledge of our intellectual and moral attributes which become sufficiently appreciable only in their collective upgrowth. After them, the true final science, i.e., morale, is able to systematize the special knowledge of our individual nature, by virtue of an appropriate combination of the two points of view, the biological and the sociological. (1852, Engl. transl., p. 357)

It took Comte nearly thirty years of constant labor to reach this position. His new advance excited him greatly.

La morale, he insisted, must deal with the *individual unity* of men. Casting around for a theory of motivation that he might employ in his new science he found only one to his taste, the phrenological system of Franz Josef Gall. What he liked about Gall's doctrine was its dynamic emphasis upon the core traits of individual men. Some men are marked by great physical courage, some by ambition, others by an altruistic sentiment. The fact that Gall tied his faculties to "interior functions of the brain" seemed to Comte to be an added scientific advantage. What Comte was seeking was a science of personality — unfortunately, some years before such a science was possible.

The science would, however, be dependent upon sound biology and sound sociology. At times la morale would lean more heavily upon its biological foundations, and would yield findings that today might be classified as physiological or as constitutional psychology. At other times, la morale would deal with the individual's place in a social and cultural context, and constitute a social psychology.

In this line of reasoning we find three striking propositions: (1) A science of individuality is possible and necessary. (2) This science, while allowing for an idiographic approach to personalities, will not shun generalizations concerning what is common to all men. It is a science of human nature in general, but one that must adapt itself to the phenomenon of individuality. (3) La morale can be viewed — as modern psychology is often viewed — either from the biological or social point of view. A few years after Comte's death both Wundt and Tarde were insisting sharply upon this same dichotomy within the science of psychology. All psychology, they said, is *either* physiological *or* social.

Comte's discovery of la morale did not weaken his enthusiasm for his earlier discovery of sociology. Sociology remains the science of the "Great Being" which encompasses individuals. The existence of a language, a culture, a social system is prior to, and in no way reducible to, the lives of individuals.

> These individual perturbations do not sensibly alter collective existence, and they do so less as it is further developed. The mutual neutralization between various individuals leaves subsisting in sociology only the permanent influences of the truly collective attributes. (1852, Engl. transl., p. 438)

Implications. The problem Comte wrestled with is perennial. The modern social psychologist, no less than Comte, is haunted by the question: How can the individual be both a cause and a consequence of society? That is to say, How can his nature depend indisputably upon the prior existence of cultural designs and upon his role in a predetermined social structure, while at the same time he is clearly a unique person, both selecting and rejecting influences from his cultural surroundings, and in turn creating new cultural forms for the guidance of future generations? And all the while this social-individual interaction is in progress, man is also, of course, a biological being subject to the laws of his species.

Whether Comte's final three sciences should stand in precisely the form in which he defined them a century ago is not here the issue. We wish merely to point to the modernity of his problem. In recent years new concepts have appeared on the horizon: Social Relations, Human Development, Behavioral Sciences, Human Relations. Each of these labels, it seems, tries to express an integration of the biological, the social, and the psychological approaches. Thus in a sense each seems to attempt a translation of Comte's untranslatable term *la morale*. While he failed completely to establish this term in the lexicon of science, he did succeed in establishing his other pivotal term, *sociology*.

Comte was aware of what he called a "happy ambiguity" in the term *la morale*. On the one side it connoted to him an abstract science of individuality, on the other hand, the realm of morality and ethics. It suggested to him the possibility of establishing a scientific basis of human morality. His view was that in order to achieve higher morality man will first have to evolve a new discipline. This science will be anchored in biological nature, and also in the study of society and culture. When its development is far enough advanced we may then hope to apply its findings for the betterment of human living, to what Comte designates "a religion of humanity." This hopeful line of reasoning has many adherents today.

To us Comte's formal and classificatory efforts may seem stilted and old-fashioned. Yet they illustrate the truth that acute problems of today were likewise acute problems in the past. And the moral is: solutions that we reach today may be sounder if we take into account solutions offered by our predecessors.

SIMPLE AND SOVEREIGN THEORIES: THEIR SIGNIFICANCE

Most social psychologists of the nineteenth century thrived on unitary explanations. Each tended to select and develop one simple and sovereign formula that seemed to him to hold the key to social behavior. The phrase "simple and sovereign" is borrowed from Henry George (1879) who declared that the single tax would be a "simple and sovereign remedy for the ills of mankind." In the nineteenth century both social scientists and reformers sought unitary solutions of social riddles. Favorite among the psychological open sesames were *pleasure-pain, egoism, sympathy, gregariousness, imitation,* and *suggestion*.

The year 1908 may be said to mark a turn away from such monistic explanations and toward pluralistic explanations. In that year the first two textbooks in social psychology appeared. The earlier, by Ross, continued to place the burden of explanation upon a single principle, sometimes called by him "imitation," sometimes "suggestion." The later text, McDougall's, called upon a plural array of instincts as prime movers of mankind. While the instinct doctrine may, in a sense, seem to represent only another simplicist type of causal force, still the instincts were regarded by McDougall as varied in type and diversified in the social behavior they provoked.

It is easy to inveigh against the fallacy of unitary explanation, or the "simplicist fallacy," as it is sometimes called. Few modern writers focus on a single motive or mechanism and claim it to be an all-sufficient explanation of social behavior. Yet even today we find authors who favor some one predominant factor to the *relative* neglect of others. Among them are such favorites as *conditioning, reinforcement, anxiety, sexuality, guilt, frustration, cognitive organization, role,* and *social class*. Thus it is difficult for even a modern author to keep a balance in his repertoire of explanations. The reason seems to be that every writer aspires to a coherent system of explanation and wishes to reduce the number of variables in his system to the minimum.

The following pages examine in some detail the leading simple and sovereign conceptions of the nineteenth century. Our purpose is not to resurrect discredited theories, nor even to warn modern authors against their own obsessive tendencies in system-building. Rather we shall hope to demonstrate that the unitary principles of the past still represent live issues in contemporary social psychology. While our present-day approach is certainly more cautious and more pluralistic, yet the phenomena described by earlier writers are still with us, and in some cases it is doubtful whether today's conceptualizations improve upon those of yesterday.

HEDONISM

We take first a simple and sovereign theory that reaches from the time of Greek philosophy into the present day. Psychological hedonism (or as Jeremy Bentham, 1748-1832, prefers to call it, the Principle of Utility) maintains that pain and pleasure are our "sovereign masters" (1789, p. 1).

While Bentham had the hedonistic doctrines of Epicurus, Aristippus, Hobbes, Adam Smith, and many others to draw on, the credit for the first extensive formulation of hedonism in relation to social psychology is his. He denies the importance of sympathy in social life. The basic fact is that men simply act to secure pleasure (sometimes called happiness) and to avoid pain. On the occasion of every act a person is led to pursue that line of conduct which, according to his view of the case taken by him at the moment, will be in the highest degree contributory to his own happiness. John Stuart Mill (1863) agreed with Bentham that to desire anything except in proportion as the idea of it is pleasant is a physical impossibility.

Starting with this basic assumption, Bentham evolves his "hedonistic calculus." It maintains that the degree of pleasure and of pain can be measured if we analyze any given affective state with the aid of certain dimensions. Pleasures (and pains) have differing degrees of *duration, intensity, certainty, propinquity* (or remoteness), *fecundity* (will pleasure or pain follow?), *purity* (will the pleasure be mixed with pain?), *extent* (will other people be involved pleasurably or painfully?). Such an analytic scheme reminds us of Wundt's and Titchener's later attempts to establish the attributes of feelings, though the dimensions they chose were different from Bentham's.

To Bentham all pleasures were measurable. Their motivational power, also their utility, could be computed. He did not deny that pleasures had differing qualities; there were, he admitted, pleasures of sense, wealth, skill, amity, good name, power, piety, benevolence, malevolence, memory, imagination, expectation, association, and relief (1789, Chapter 5). But one was "as good as" another. A bottle of pickles is as good as a book of poetry if it yields equal pleasure. (On this point John Stuart Mill disagreed.) All that mattered to Bentham was the maximizing of total pleasure. To reach a rational decision regarding a course of conduct — and psychological hedonism has a heavily rationalistic cast — one would simply compute the probable hedonic consequences of the proposed act in terms of the resulting duration, intensity, certainty, etc., of pleasure. He would (and should) act so as to bring himself the greatest hedonic good.

At this point psychological hedonism merges into ethical hedonism. Not only is it true that men do tend to act in such a way as to maximize pleasure but they *should* so act. To be sure, other people's pleasure may enter into the person's calculations — in the dimension of "extent." He should try to maximize the happiness of others along with his own. And the policy of government and of all social agencies should be to facilitate "the greatest good (happiness) to the greatest number."

Thus utilitarian ethics is simply an extension of a psychological theory into the domain of morals. To apply this ethics one must in general follow a policy of laissez-faire, or "hands off." It was customary to ask the question "How much government should we have?" and to give the answer "As little as possible." For if men instinctively seek pleasure, then they need very little guidance or constraint. Let each serve his own good. While an occasional law may be needed, still laissez-faire is the surest way to achieve the greatest good for the greatest number.

Here is a case where a psychological assumption had far-reaching practical consequences in society. Its proponents — and there were many — fought against social legislation. Why, they asked, should we curtail the good resulting from the industrial revolution? For example, if under sweat-shop conditions ten thousand men obtain factory-made pants at a cheaper price, the "greatest good" is served; for their pleasure in the aggregate probably exceeds the pain caused to the relatively few sweat-shop (or child) laborers involved in the process of mass production.

The doctrine of the "economic man" is likewise a direct outgrowth of the same hedonistic assumption. Man, it was held, basically works for monetary gain, because money is pleasure *in abstracto*. With it one buys the utilities that yield him maximum good according to his conception of the matter (and no one else's conception counts). To take only one example from the classic economic laws founded on this assumption, it was held that "man will buy in the lowest market and sell in the highest market." For many years this "law" of behavior underlay economic theory. To a considerable extent it still does so. Yet exceptions are numerous. In times of panic, the reverse relation holds. Further, a man who is temperamentally generous or temperamentally pessimistic may act contrary to the law. And Pareto tells of a superstitious banker who lost money by refusing to sit at a table with thirteen people to consummate a sale "in the highest market." Thus a host of noneconomic motives enter into economic conduct. It is true that many modern economists have abandoned the rigid formulas that derived from rationalistic hedonism. The "economic man" is becoming extinct. A newer, pluralistic conception of economic behavior is arising that takes into account many more factors than are included in the hedonistic assumption (Parsons, 1949; Katona, 1951).

Herbert Spencer, 1820-1903, gave support to the hedonistic theory and to laissez-faire by tying both to the doctrine of evolution. He pointed out that pleasurable activities, for the most part, make for survival. Pain betokens danger and death. It is therefore natural to seek pleasure and avoid pain. To the obvious objection that some destructive activities are also pleasurable (drunkenness and other vices) Spencer replied that this dislocation is only a reflection of the imperfect state of society. When social life is fully regulated by evolutionary principles, then pleasure and survival value will be identical.

Spencer joined the opponents of social legislation, likewise on psychological grounds. The state must respect rugged individualism and allow every pleasure-seeker to attain his own good in his own way — so far as possible. Men's motives, like nature herself, are often "red in tooth and claw," but even so a premium must be put on these natural impulses. To be sure, Spencer also recognized the affiliative (sympathetic) motives in men. But these, he thought, should be expressed not in the activities of the state, but only within the circle of the family. For him there was a sharp distinction between the "ethics of the state" and the "ethics of the family." On this ground he opposed the advent of free public education in Britain — which was finally achieved over his protests in 1873. Let the family plan for, and pay for, the education of its children. The state should never put a premium on weakness.

We have dwelt on the relationship between psychological hedonism and social policy in the nineteenth century for two special reasons. First, we wish to show that here is a psychological hypothesis that was, in a sense, put to the test — and failed. In the long run the state could not leave men alone to seek their own good, relying on their instinct for pleasure and their reason. The cry of protest from Dickens, Ruskin, Kingsley, and the socialists finally prevailed. Their argument was that only one side of human nature and only one portion of the population were being favored by Utilitarian ethics. Human life is too complex to rely on the manifestly narrow doctrine of Utility. The tide turned toward social welfare policies — toward what today is called "welfare state." The claims of simon-pure hedonism, when tested in the crucible of social policy, proved inadequate.

A second reason for dwelling on this matter is to provide illustration for another significant fact. The theories of social psychology are rarely, if ever, chaste scientific productions. They usually gear into the prevailing political and social atmosphere. Dewey (1899) has shown that this is so. For example, an aristocracy produces no psychology of individual differences, for the individual is unimportant unless he happens to belong to the higher classes. Dewey likewise points out that dualistic psychology flourishes best when one group holds a monopoly of social power and wishes to do the thinking and planning, while others remain the docile, unthinking instruments of execution. And apologists for the status quo, he adds in another context, are the ones who most readily declare human nature to be unalterable.

> The ultimate refuge of the standpatter in every field, education, religion, politics, industrial and domestic life, has been the notion of an alleged fixed structure of mind. (1917, p. 273)

Our point is that laissez-faire was in part the consequence of a psychological theory, and in part a product of the prevailing social ethos. The thinking of Bentham, Mill, Spencer, and other hedonists was inevitably molded by the prevailing practices of the era of the industrial revolution. Their psychological theory meshed into the social situation of the day, and became to some extent what Marx and Engels (1846) and Mannheim (1936) would call "ideology."

In recent years we have had more spectacular samples of the dependence of scientific theory upon prevailing political and social conditions, viz., in the surrender of German psychology to Nazi thought and of the Russian psychology to Communist thought.

Hedonism today. Yet hedonism is more than a culture-bound and time-linked ideology. It is a basic theory of the prime motivation of men. In various guises it is held today as tenaciously as in the past. Freud, especially in his earlier writing, insisted that all instincts "seek pleasure." While they may be overlaid, repressed, or controlled, they still strive unconsciously, and in disguised ways, for pleasure.

At the present time, however, the term "pleasure" is not popular with psychologists, chiefly because it implies a dualism in the relation of body and mind. (Pleasure is a conscious state, and it is considered unscientific to claim that consciousness "causes" behavior.) Hence other labels for current hedonism have been devised. For example, Dollard and Miller write:

> The principle of reinforcement has been substituted for Freud's pleasure principle. The concept of "pleasure" has proved a difficult and slippery notion in the history of psychology. (1950, p. 9)

As we shall see, "reinforcement" refers to past pleasure (or satisfaction) rather than to future pleasure (which the classical hedonists stressed), but the hedonistic flavor is still present.

Another disguised version of hedonism is the prevailing dictum, found in most present-day psychological texts, that all motives tend to achieve "tension reduction." Thus Kluckhohn and Murray write that "the nearest thing to an all-embracing principle" is the concept of "need, drive, or vectorial force." Such a force leads the individual to "re-establish equilibrium, reduce the tensions, appease the need" (1948, p. 14). This view is at bottom a variant of hedonism, because the act of escaping tension is also an escape from anxiety, from frustration, from distention — in brief, from pain. To be sure, the stress here is on negative hedonism (Epicurean) rather than a positive pleasure-seeking hedonism (Cyrenaic), but it is hedonism all the same. Since *all* motives are said to be of this tension-reducing order, the position is not unlike Bentham's contention that "pleasure and pain are our sovereign masters."

Hedonism, therefore, as an over-all conceptualization of human motivation, is neither dead nor discredited. Perhaps its longevity certifies that it contains a core of truth.

Critique. There are, however, many arguments raised against hedonism, some of which apply to modern as well as to classical formulations. To assist our survey of these criticisms, we shall adopt the analytical scheme proposed by Troland (1928). This author advises us to distinguish between hedonism of the *present,* of the *future,* and of the *past.*

Hedonism of the *present* would hold that each individual now, in the present moment, is acting in such a way as to obtain maximum present pleasure. This is clearly an untenable position, excepting perhaps in the case of a vegetative, psychotic patient. Most people do submit themselves to the dentist chair; they do wash dishes, save money, and forego present pleasures. No hedonistic writer makes any contrary claim.

Hedonism of the *future,* however, is the doctrine of Bentham and Mill and, in some respects, of the "tension reduction" school of thought. The individual is doing now what in the future is designed to gain happiness or relief from pain. Here many criticisms are relevant.

(1) We recall that when recast in terms of social policy (laissez-faire) the assumptions of classical hedonism failed to hold up. Men did not in fact act so as to maximize pleasure for themselves or for others.

(2) Troublesome too is the interactionist position implied. How can "ideas" of future pleasure motivate present conduct?

(3) Especially serious is the charge that men cannot aim at pleasure in the abstract. They can only seek concrete goals (a job, a mate, food, shelter). At best the feeling-tone is a by-product of reaching the goal; it does not constitute the motive itself. In this connection, employing the evidence of introspection, Titchener finds that pleasure and pain are not the sole determinants of action. Indeed, he writes, "I do not consider that they can be numbered at all among the conditions of action" (1908, p. 297f.).

(4) Similar is McDougall's argument that pleasure and pain are nonmotivational; they are nothing more than signposts, guides, indicators that instincts are successfully or unsuccessfully running their course (1923). A mother cannot help but care for her offspring through thick and thin (if she has a normal parental instinct). Pleasure and unpleasure merely tell her that the instinct is, or is not, fulfilling its basic purpose of successfully caring for her young. Instincts are the prime movers; at most they are guided by the signals of pleasure and unpleasure to select instrumental activities that will fulfill their purposes.

Freud, we recall, identified instincts with pleasure seeking. His earlier view of the matter was thus very different from McDougall's. In later years, however, Freud modified his hedonistic position considerably, and in *Beyond the Pleasure Principle* (1920) postulated Thanatos, a death instinct, which ultimately negates Eros and pleasure seeking.

(5) If we appeal to common observation, doubts increase. While hedonism may seem a good "fit" to the self-seeking activities of childhood and youth, it seems less appropriate to the activities of parents, teachers, artists, and above all, martyrs. In mature conduct there is often an overwhelming sense of duty, loyalty, or commitment to a way of life, and these motives seem to take control of the personality. When pleasure or pain are encountered, they may be regarded as "twin impostors," and the

motive may run its course regardless of the affect involved. To this line of argument the hedonists reply that duty-ridden people are after all merely avoiding the pain of not doing their duty; and martyrs too are selecting the least painful course of conduct, or perhaps seeking the pleasure of the next world. It is a question whether at this point hedonists are not overextending their terms. Can the bomber who crashdives into his target be said, in any reasonable sense, to be seeking pleasure and avoiding pain? His motive seems remote from such affective considerations. It is easy to stretch the meanings of pleasure and pain until they lose significance.

Hedonism of the *past* escapes some of these criticisms, but is open to others. Its most succinct statement comes from Thorndike's dictum that "pleasure stamps in; pain stamps out" (1898). He was referring, of course, to the learning process, and defining what he called "the law of effect," which in the present day is usually rendered as "reinforcement theory." This form of hedonism holds merely that each person now is acting according to the mode of conduct he has in the past found pleasurable (satisfying, effective, tension-reducing). Good consequences have reinforced his habits, his opinions, his ways of seeking goals, so that now he maintains them. To argue this position in detail would take us too far afield. Suffice it to say that studies of memory fail to show convincingly that pleasurable memories are significantly more often retained than unpleasurable, as the theory would predict (Meltzer, 1930). And the existence of satiation, boredom, risk-taking, and other phenomena make trouble for the theory (Hoppe, 1930; Allport, 1946).

At the present time no conclusion seems justified, other than that pleasure and unpleasure have some important relationship to motivation. Just what it is we do not yet know. It may be as central as Bentham claimed, or as incidental as McDougall claimed. It may be a principle appropriate to childhood but less so to maturity. In any case, the role of affect needs vastly more empirical and theoretical study before it finds its rightful place in an improved science of social behavior.

EGOISM (POWER)

Hedonism is a doctrine of self-centeredness. A person who seeks pleasure and avoids pain is inevitably serving his own affective interest. Even though, as Mill eloquently argued, he can sometimes gain happiness from contributing to the welfare of others, it is still the pursuit of his own pleasure that motivates him.

Thomas Hobbes, 1588-1679, was likewise a hedonist, but unlike Bentham and Mill he reduced pleasure seeking to a more basic passion of egoism. Egoistic conflict was, to his mind, so prominent that mankind must be regarded as in a state of "war of all against all." To obtain pleasure it is necessary first and foremost to have power. For power gives one the means to obtain ease and sensual pleasure; it brings admiration and flattery from others, and it thus conduces to "the greatest joy of the human soul," which is to have a high opinion of oneself. Therefore, man's most basic motive is "the desire of power after power that ceaseth only in death."

> So that in the first place, I put for a generall inclination of all mankind, a perpetuall and restlesse desire of Power after power, that ceaseth only in Death. And the cause of this, is not alwayes that a man hopes for a more intensive delight, than he has already attained to; or that he cannot be content with a moderate power: but because he cannot assure the power and means to live well which he hath present, without the acquisition of more. (1651, p. 63)

Since life in society would be impossible if the insatiable desire for power were not curbed, men yield to the "common power" of the state which he calls the *Leviathan.* Through a form of "social contract" each person obtains protection from other power-seeking mortals, and obtains a chance to utilize his leisure for the arts of peace and for the acquisition of knowledge.

In several respects Hobbes gives an analysis of the power motive that is surprisingly modern. The following passage indicates the extent to which power-holders may corrupt communication and create an ideology favorable to themselves:

> For I doubt not, but if it had been a thing contrary to any man's right of dominion, that the three Angles of a Triangle should be equall to two Angles of a Square; that doctrine should have been, if not disputed, yet by the burning of all books of Geometry, suppressed, as farre as he whom it concerned was able. (p. 67)

Extravagant as the example may seem, its resemblance to the deeds of Nazis and Soviets is obvious.

Equally shrewd is his observation of the way in which each man thinks very highly of his own particular ego.

> For such is the nature of men, that howsoever they may acknowledge many others to be more witty, or more eloquent, or more learned; Yet they

> will hardly believe there be many so wise as themselves. (p. 82)

Since all men have a maximum feeling of self-worth and self-satisfaction, then men are by nature equal. "For there is not ordinarily a greater signe of the equall distribution of any thing, than that every man is contented with his share." In many respects Hobbes foreshadows modern doctrines of self-esteem, status-seeking, and self-regard as pivotal motives, and at the same time perceives the sly operation of what today are called ego-defensive mechanisms.

Many variants of the power doctrine can be found in the nineteenth and twentieth centuries. In Germany two influential proponents were Max Stirner and Nietzsche. Stirner's book, *The Ego and his Own* (1845), anticipated the works of Nietzsche. The latter in his *The Will to Power* (1912), gives a crisp statement of his position:

> What man wills, what every smallest part of a living organism wills, is a plus of power. Both pleasure and unpleasure are consequences of striving for it. (Section 702)

To Nietzsche, as to McDougall, hedonic states were incidental to a more basic form of motivation. All social behavior, according to Nietzsche, is a direct or disguised reflection of power seeking. Affect is a "consequence." Even the emotion of gratitude is nothing more than a "good revenge" one takes to re-establish his own potency in a social relationship. When viewed in the light of man's will to power, his alleged motives of altruism, love of truth, and religion turn out to be hollow shams.

A generation later the doctrine received an empirical and clinical rendering at the hands of the Austrian psychiatrist Alfred Adler (1917), who felt that his teacher and associate, Freud, was neglecting the role of inferiority feelings so often engendered by poor physique or organic defects. The striving for masculinity, overcompensation, a desire "to confirm one's own value," loomed large in Adler's thought. Still later another psychoanalyst, deviating from Freud, declared competition to be even more important than sex striving in causing personality disorders (Horney, 1939), and the philosopher Bertrand Russell (1938) emphasizes the same motive. While present-day psychologists avoid a unitary conceptualization, much current work features the status motive (Williams, 1920), the ego-level (Hoppe, 1930), ego-involvement (Allport, 1943), self-image (Lecky, 1945). These, and many other writings, lay heavy emphasis upon self-esteem.

The most uncompromising and exhaustive treatment of the ego-motive is found in *L'Égoisme,* written by the French biologist, Felix Le Dantec (1918). His approach is evolutionary. Le Dantec argues that fairly stable levels of matter and life have evolved: atoms, molecules, cells, organisms. Each tenaciously seeks to maintain itself. A phagocyte in the blood stream does not devour disease bacteria "in order to" serve the organism as a whole, but simply to maintain itself. Similarly, each individual organism pursues its own good, without reference to society. Try as hard as we may, we cannot evolve a social unit to supplant the individual ego. Each person is interested exclusively in himself; altruism is an illusion. "To exist at all is to struggle; to survive is to conquer." All social varnish is superficial; at bottom the caveman remains. Even the most pious Christian is thrilled and excited by battle and by brute victory.

If this stark account of human motivation is true, one naturally asks, "How are we to account for society at all, especially for the social institutions that seem to foster cooperation and mutual aid?" Le Dantec answers: it is because of the existence of common enemies. The basis of all society is *l'ennemi commun.* Even the family is at bottom merely a serviceable unit for common protection. Parents take care of children in order later to benefit, e.g., gaining financial support in their old age. The basis of friendships and of in-groups is not love so much as a common fear or common hate. The easiest way to get along with a stranger is to find some person, some group, or some object that both dislike. Allies in wartime place their own selfish interests in abeyance until the common enemy is defeated; then the allies fall out with one another. Indeed, even a single nation has genuine unity only in wartime, when the lives of all its members are threatened. For this reason governments provoke wars.

Yet, Le Dantec continues, we cannot afford to admit that our motives are purely egoistic and that the only basis of social solidarity is common hate and distrust. To do so would weaken further the already precarious fabric of society. Hence a fantastic superstructure of "metaphysical absolutes" has grown up. A metaphysical absolute is any doctrine, taught with the sanctity of tradition, that masks the basic egoism of human life. Conscience is the carrier of such doctrines, and is created by parents and teachers in children in order to preserve the hypocrisies that are so essential in social life.

Take, for example, the metaphysical absolutes of religion. The Decalogue, so it seems to Le

Dantec, provides a transparent example of social hypocrisy. The first three Commandments require men's obedience to God, and this initial obedience provides the sanction for all remaining Commandments. Le Dantec would agree with Voltaire, who held that no parent could bring up a child without invoking divine sanction. The postulated power of a deity is essential to achieve social control. The fourth Commandment reveals the point: it says "Honor thy father and thy mother" — a rule of extreme convenience to parents. The utilitarian nature of this Commandment is betrayed in its additional phrase "in order that thy days may be long . . ."

"Thou shalt not kill" is likewise an egoistic rule, meaning, naturally, that thou shalt not kill *me*. Similarly "thou shalt not steal" protects the property of the one who teaches the rule. "Thou shalt not bear false witness" grew out of the necessity for the in-group to be relatively truthful. Likewise, the prohibition against coveting is psychologically adroit, for the place to block transgressions is in the attitudes of the would-be transgressor before they issue into action. "Thou shalt not commit adultery" is a commandment taught with special enthusiasm by those who have sexual partners to those who have not.

Le Dantec's view of human motivation and of social values may seem to an extreme degree to be cynical and dyspeptic. But is it wrong?

Critique. Common to all theories of egoism is one precarious assumption — namely, that the process of social learning never results in a genuine transformation of motives; that is to say, the process of socialization is never truly effective. Le Dantec views socialization as a matter of "deformations of the ego resulting from life in common." "Why deformations?" one may ask. May not social learning result in formation, in conformation, or even in re-formation? Comte held a more optimistic view of socialization, expressed in his Law of Affective Evolution, which says that with time there comes a diminution in the preponderance and intensity of personal inclinations, and a genuine growth and extension of altruistic sentiments.

May it not be that the picture of egoism is appropriate to the unsocialized being (including the child) more than to adults with socialized personalities? Hobbes wrote, "the wicked man is the child grown strong." Unbridled self-assertion may be natural at an early stage in the development of personality, but abnormal at a later stage. Theories of socialization today are split on this same issue: some hold that the primitive egoistic drives are never modified, but only overlaid, or channelled, or "cathected." Others argue for a changing ego-structure, wherein new interests become functionally autonomous of their origins, or where productiveness and maturity effectively inhibit earlier instinctive aggression and pleasure seeking.

It is difficult to strike a proper balance in this matter. On the one hand we cannot deny that hypocrisies and rationalizations of men concerning their own motives are common. On the other hand, who can say with certainty that the original nature of man does not include some strongly affiliative dispositions? Or that the course of social learning does not result in thoroughgoing transformations of the character-structure? While self-love remains positive and active in every man, is it necessary that it should be sovereign with him? We shall resume consideration of these questions in our discussion of Sympathy (pp. 18-21).

Irrationalism and Rationalism in Social Psychology

In this section we shall interrupt our exposition of simple and sovereign doctrines to discuss an issue that is central to all of them. The reader has noted, no doubt, that hedonism and egoism are regarded as powerful and irrational pushes that dominate man's entire behavior, making life in society at best a difficult achievement. This unfavorable view of man's social capacities has in fact predominated throughout the ages. From the time of antiquity writers have stressed the force of pleasure seeking, of fear, of the power drive, or other equally blind passions. In *The Prince* (completed in 1513) Machiavelli saw nothing to be gained in appealing to people's intellect. The prince must see to it that he is beloved and feared. If he must choose one of the two, let him inspire fear, for people are at bottom ungrateful, false, cowardly, and interested only in their own welfare. A century and a half later Hobbes, as we have seen, held an equally uncomplimentary view of human nature.

Relative rationalism. Yet as the Enlightenment broke over Europe the stock of the common man went up. He was given credit for managing his social life with relative (though not complete) intellectual acumen. Locke, 1632-1704, felt that the "social contract" was a rational necessity of man's nature. In fact, he equated "natural law" with "the command of reason." Unlike Hobbes, who thought the state of nature was a state of war, and unlike Rousseau, who thought it was a state of happy innocence, Locke saw the state of nature to be one of reasonableness, moderation, adaptive compromise (1689). The doctrine of the reasonable decency of man, coupled with faith in his ability to utilize the developments of science, commerce, and industry for his own ends was spread by such writers as Godwin (1793) and Condorcet (1795), and influenced both the Utilitarians and the founders of American democracy.

At first sight it seems paradoxical to call hedonism a rationalistic type of social psychology. The aim of men is emotionally determined; they cannot abstain from the pursuit of pleasure. But when this starting point is admitted, then their subsequent conduct is accomplished through rational choice—perhaps even through Bentham's "hedonistic calculus." Present-day forms of hedonism have completely lost their rationalistic cast, but Bentham and Mill had high regard for man's capacity to calculate in advance the consequences of his conduct.

In the United States the Founding Fathers likewise had marked faith in men's ability to face facts, judge rightly, and vote wisely. The conviction, to be sure, was stronger in Thomas Jefferson and James Madison than in Alexander Hamilton, who feared, as did Plato, that democracy runs the danger of becoming the "rule of the worst." In the *Federalist Papers* the issue is argued out. Checks and balances are devised for the Constitution which, while maintaining faith in the rational judgment of voters, yet prevent government by mob passions or by power-hungry demagogues. Social psychologists today might well ask themselves whether the issue of rationality versus irrationality did not achieve a more balanced solution in various historic American documents than it achieves in their own theories. After all, American democracy has proved viable for nearly two centuries. Its assumptions concerning "the nature of the men governed" cannot be wholly wrong. Therefore we may well search *The Federalist* (1788) and other writings of Jefferson, Madison, and their contemporaries for a well-balanced view of rational and irrational factors in the social nature of man.

Nineteenth-century irrationalism. While the rationalism of the Enlightenment was not without its influence, it remains true that most social psychologists in both past and present have favored irrational interpretations. A theory is irrationalistic if it possesses two features: (1) It must claim that the prime determinants of human behavior are emotion, drive, instinct, or some form of blind will. (2) It must also square these motives with the apparently peaceful and logical structure of group living. More specifically, it must tell what the function of intellect is. If the intellect has no causal properties, then it must at the very least be granted a role in inventing hypocrisies and rationalizations that help sustain the social edifice.

An influential expression of irrationalism in the nineteenth century is Schopenhauer's essay, *The Primacy of the Will in Self-Consciousness* (1819). In it he gives many reasons for believing that blind unrest is basic in all life. When the will (emotional need) is active the intellect can only serve it. A more modern statement of the same position is found in the writing of Kempf (1918) who holds that the autonomic nervous system is the leader and the central nervous system is the servant, in their co-functioning.

It was the gigantic impact of Darwin, McDougall, and Freud, with their varying conceptions of instinct, that assured the supremacy of irrationalism in present-day social psychology. The depreciation of intellect was further assisted by the rise of behaviorism, and by the union of social psychology with abnormal psychology during the late ninetenth century. Only in recent years, as we shall soon see, has there been any marked sign of reaction against irrationalism.

Rationalization. Irrationalists admit that for all our enslavement to instincts, to blind will, or to the push of environmental stimuli, men do perceive; they do cognize, plan, and ponder. And in particular they perceive, plan for, and conceptualize their social relationships. Irrationalism insists, however, that much of man's theorizing is merely a mask for the real state of affairs, specifically for the unconscious motives that are operating. Man's hypocrisies were clearly noted by Machiavelli, by Hobbes, and many other earlier writers. Bentham called them "fictions," which he defined as the imaginary entities people put in place of perceptible and inferential entities. Fictions have the happy gift of smoothing the course of social life. For example, instead of claiming his prerogatives as a powerful individual, the king resorts to double-talk about "the crown," thus masking his own selfish demands. Bentham introduced likewise the concept of "fallacy," by which he sought to label all misuse of praise and blame in social control. Today we speak of "name-calling," of "smear words," and "purr words." This trick Bentham clearly spotted and named, a century before the days of modern semantics (Ogden, 1932).

The term *rationalization* was first used by Ernest Jones (1908) to represent Freud's finding that irrational acts are "justified by distorting the mental processes concerned and providing a false explanation that has a plausible ring of rationality." A few years later Robinson in his *Mind in the Making* (1921) applied the concept on a large scale to human history, defining it as follows:

> Rationalization is the self-exculpation which occurs when we feel ourselves or our group accused of misapprehension or error. (p. 44)

Actually the term has broader meanings, reaching beyond sheer self-justification. It refers also to the cloaking of a motive, to defense of a value system, to excusing the bad behavior of our loved ones, and to any refusal to look at an argument in its own right.

Closely related is the doctrine of *stereotype* advanced by Walter Lippmann (1922). The "picture in our heads" serves a variety of useful functions because it simplifies complex social reality to manageable, if false, proportions. The dynamics of the stereotype are clearly set forth in the following quotation.

> If the experience contradicts the stereotype, one of two things happens. If the man is no longer plastic, or if some powerful interest makes it highly inconvenient to rearrange his stereotypes, he pooh-poohs the contradiction as an exception that proves the rule, discredits the witness, finds a flaw somewhere, and manages to forget it. But if he is still curious and open-minded, the novelty is taken into the picture, and allowed to modify it. Sometimes, if the incident is striking enough, and if he has felt a general discomfort with his established scheme, he may be shaken to such an extent as to distrust all accepted ways of looking at life, and to expect that normally a thing will not be what it is generally supposed to be. (p. 100)

Much older, but closely related, is the doctrine of *ideology*. Some trace it to Francis Bacon (1620) who pointed to the chronic error and distortion that lay in the prevailing "idols" of the tribe, of the marketplace, and of the cave. Marx and Engels (1846) used the term "ideology" to refer specifically to the elaborate beliefs promulgated by the capitalist class to justify its favored position in society. Influenced by Marx and Engels, the movement known as the "sociology of knowledge" (Mannheim, 1936) has concerned itself with any ideology that tended to derive from, and to justify, a set of values that was in fact determined by irrational, or quasi-rational, considerations. It is interesting to note that the present-day usage of the term is losing the savor of hypocrisy. Whereas Marx would never have called communism an "ideology," today it is customary to do so. We speak even of democratic or Christian ideology without derogatory implication.

Among other variants of rationalization we may cite Le Dantec's *metaphysical absolutes*, previously discussed; *myth* as used by Sorel (1914) and Delaisi (1927); *folklore* (Arnold, 1937); *vital lies* (Lee, 1912); *"thobbing"* (Ward, 1926); and *intensional orientation* (Hayakawa, 1941).

Special attention should be called to Pareto's doctrine of *derivations*. Pareto was a versatile economist who wrote a vast treatise in sociology from the point of view of an irrationalist social psychology (1916). Apparently Pareto was unacquainted with the writing of his intellectual predecessors. The results was therefore a *de novo* system of thought. Those who recognized that he was only saying, in different and complicated terms, what irrationalistic social psychologists had previously said severely criticized him (e.g., McDougall 1935); but those who, like Pareto himself, were unacquainted with the history of social psychology, praised him extravagantly, likening his methodology to Descartes' and his originality to Aristotle's (Bousquet 1928).

Pareto, like all irrationalists, offers a theory of the real motives of men (*residues*) and of fictitious motives (*derivations*). Residues bring men under the domination of "sentiments" which clever people manipulate by the use of "nonlogical" appeals. The resulting social conduct is "explained" and prettified with the aid of verbal illusions (derivations). Derivations spread widely in society through devices that psychologists today recognize as principles of propaganda. The true motives lie in classes of mental processes to which Pareto gives such labels as "the instinct of combinations," "need of manifesting one's feelings by exterior acts," "sexual residues," and the like. We need not take time here to show how both doctrines (residues and derivations) are handled in more orderly fashion by various other systems of irrationalist psychology (e.g., Freudian, McDougallian).

Twentieth-century rationalism. While the dominant trend in social psychology has certainly been irrationalistic, counterinfluences have not been lacking. One might cite, for example, the general cast of John Dewey's thought. At one time, to be sure, he argued that "instinct" underlay all social conduct (1917), but he soon shifted his formulation to favor "habit" as the unit of behavior (1922). Habit in Dewey's hands is not wholly irrational, rather it is a learned disposition that aids man in his rational transactions with his environment. While Dewey is by no means blind to the force of emotions, his tenacious faith in education and in democracy has, like the teaching of Thomas Jefferson, acted as a brake upon irrationalism in American social psychology.

In this connection we should mention the invention and growth of opinion polling within the past fifteen years. The movement assumes that the conscious views and judgments of men have extensive influence on their conduct, and should play a part in the determination of public policy. The widespread use of the polling method in the United States Army during World War II (Stouffer *et al.,* 1949) reflected far more rationalistic than irrationalistic assumptions concerning the social nature of man.

We have said that McDougall's instinct doctrine (1908) augmented Darwinian and Freudian influence. Almost immediately it found its way, via Graham Wallas (1908), into an interpretation of political life. A few years later, however, Wallas became alarmed at the one-sidedness of the resulting picture, and wrote a second treatise to counteract the irrationalism of his own first production. In *The Great Society* (1914) he proposed in addition to the McDougallian list of native propensities an "instinct of thought." The greatest triumphs of human history, he argued, were products of this instinct, unaffected by the turmoil engendered by all other instincts. The instinct of thought is the only instinct aroused by inanimate objects, and thus productive curiosity is a genuine basis for man's transactions with the physical portion of his environment. Nor is it entirely absent in social relationships.

While it would take too long to argue the matter here, it seems safe to say that post-Freudian psychoanalysis has taken a certain turn toward rationalism. The crux of the change lies in a revised view of the ego. Whereas Freud declared that the ego (by which he meant the rational region of personality) had no dynamic power in its own right, neo-Freudians perceive that the ego may genuinely transform the instinctual forces of the id until a "productive" (more rational) personality results (*cf.* G. W. Allport, 1953). Equally striking is the strong drift of favor in psychotherapy toward "nondirective counseling" (Rogers, 1942). This movement reflects confidence in the rational powers of the self to survey its situation and to reach a reasonable decision regarding its future course of conduct.

Finally, among the rationalistic trends we mention the growing emphasis on phenomenology and on cognitive organization (e.g., MacLeod, 1951; Krech and Crutchfield, 1948). While this movement does not deny the role of motive and emotion, it prefers to look at these factors not as the starting point of all action, but as factors present within the total perspectival grasp with which the individual views his situation. Surveying the irrationalism of earlier social psychology, Asch has declared that it has drawn a caricature rather than a portrait of man (1952, p. 24). Asch argues that people ordinarily make use of their capacity to grasp accurately the objective character of a social situation, and to act in keeping with its rational requirements. Phenomenology and Gestalt psychology center attention upon men's perception and cognition of their surroundings. This approach, as contrasted with the approach through motivation and urge, will capture such rational factors in social conduct as may exist.

SYMPATHY

It is not only rationalists who object to hedonism and egoism as sole explanations of social conduct. Some irrationalists (believers in instinct, urge, inborn nature) join the protest. They argue that self-love is a grossly one-sided principle. Granted that self-love is present in every man, they say it is not necessarily dominant in his nature.

Some even reverse the formula; they insist that the human relationship of love or symbiosis is primary; e.g., Kropotkin (1902), Suttie (1935), Ashley-Montagu (1950). A basic groundwork of love and trust is established in the early mother-child relationship; it is essential to all survival. So fundamental is this fact that it is often overlooked by psychologists, who are inclined to glue their attention upon the pattern of hostility and aggression which is perceptible only because it contrasts with the underlying ground of affiliation (see Sorokin, 1950; especially Chapter 5).

Other theorists do not insist that sympathy or affiliation is sovereign, but they agree in giving it a prominent place among social motives. In this section we shall consider the history of this motive whether or not it is advanced as a sole explanation of social conduct.

SYMPATHY, IMITATION, SUGGESTION

First let us note that there is good historical reason why we should examine three fundamental concepts in close succession. In a sense, they compose the principal triumvirate of theories in social psychology. The reason for their close relationship goes back to Plato. As everyone knows, Plato, like Freud, conceived the human mind as made up of three faculties or "institutions." To Plato the abdomen was the seat of emotions or feeling; the breast the seat of striving and action; the head the seat of reason and thought. (Society had three parallel classes: slaves, warriors, philosophers.) Throughout the ages this Platonic trichotomy has persisted. Mind, it is said, is constituted of:

Affection (feeling)
Conation (striving)
Cognition (thought)

Now it is possible for a psychologist to stress one of these faculties at the expense of others. When a social psychologist does so, seeking in any one of these functions an exclusive (or almost exclusive) explanation of social conduct, he is likely to end up with a simple and sovereign system based on one of three principles:

Sympathy (affective)
Imitation (conative)
Suggestion (cognitive)

In a certain sense these concepts are interchangeable, in that the phenomena one author would explain by one principle, a different author would explain by another. And occasionally a given author combines two, as when Ross (1908) speaks of the "suggestion-imitation" principle.

Because of this historical linkage, we shall consider the three concepts *seriatim,* starting with the doctrine of sympathy in its various forms.

PROPONENTS OF SYMPATHY

Adam Smith, in his *Wealth of Nations* (1776), gave marked impetus to hedonism and to laissez-faire. Yet his total system of thought demanded an equal emphasis upon human sympathy. The opening sentence of his *Theory of Moral Sentiments* (1759) reads:

> How selfish soever man may be supposed, there are evidently some principles in his nature which interest him in the fortune of others, and render their happiness necessary to him. . . .

Smith distinguished two basic forms of sympathy which most subsequent writers have preserved. First there is the quick, almost reflex, type of response. When we see a person struck with a stick we cringe; when we watch a tightrope walker we grow tense. In these cases we feel as the other person feels and do as he does. Some writers have regarded this mimicry as instinctive. But today we might say that the conditioned reflex formula covers it well enough: at first we cringe when we are hit, and later we cringe at the visual cues that were initially associated with our own cringing.

The second type of sympathy is more intellectualized. We are able to sympathize with a person even though we do not feel as he feels. We may congratulate the successful and condole the afflicted at times when we feel neither success nor affliction. There are sympathetic bonds between friends even when they are viewing a given situation with unlike emotions.

Smith derives man's sense of justice from the operation of sympathy. Our capacity for sympathy leads us to feel with both the aggressor and the aggrieved. The resulting balance of sympathy determines our final judgment as to where justice lies. Similarly, our judgments of propriety or good taste are the result of weighing a person's emotional expression against the exciting fact. If it seems disproportionate we cannot sympathize, and so judge his response to be improper. These instances indicate how Smith uses his favorite principle to explain a large and diversified range of social behavior.

Herbert Spencer distinguishes the same two basic forms of sympathy (1870). He called them respectively *presentative* (immediate, reflexive) and *representative* (conscious, reflective). For a still more highly intellectualized sentiment (e.g., the abstract idea of Love as a proper way of life) he coins the term *"re-representative* sympathy." Also, like Smith, he felt that the hedonism and laissez-faire he espoused needed softening with the aid of a principle that would explain affiliative urges. Society originates in the sex instinct. It leads to the institution of the family, which is the essential unit of survival. Within the family the principle of sympathy takes over; a child survives not because he is strong and fit but because his defenselessness arouses some grade of sympathy in others. Spencer, as we have previously seen, argued strongly that this instinct should find its expression only in the family circle and not be allowed to weaken the rugged fiber of the state, which should give its rewards to the strong and never to the weak. Gregarious propensities are a mutation in the evolutionary struggle. While they are essential to human survival, they must not lead to coddling. Spencer, as we have said, had no use for the welfare state; he even opposed free public schools. Let education, he said, be a product of family sympathy, not of public policy.

Peter Kropotkin. The harshness of the "red in tooth and claw" doctrine of evolution led to a marked reaction. In *Mutual Aid* (1902) Kropotkin argued that the mutation of sympathy to which Spencer called attention is the *primary* fact in human evolution. It represents the neglected aspect of Darwinism. Kropotkin's evidence rests chiefly upon anecdotes collected both in time of war and in time of peace. He argues that the mutual aid we see around us is not based upon love for specific individuals, but rather upon a broad instinct of "human solidarity." A more modern book that in many respects restates Kropotkin's position is Ashley-Montagu's *On Being Human* (1950).

Gregariousness. Many authors have postulated, as Kropotkin did, a protean social instinct. Sociologists especially have done so, in order that this broad psychological foundation might support their differentiated theories of social structure and functioning. For Giddings there is a basic "consciousness of kind" compounded of organic sympathy, the perception of resemblance, conscious or reflective sympathy, affection, and the desire for recognition. "To trace the operation of the consciousness of kind through all its social manifestations is to work out a complete subjective interpretation of society" (1896, p. 19).

During the first world war Trotter published his *Instincts of the Herd in Peace and War* (1916), which postulated a similar basic explanatory propensity. In it Trotter argued somewhat chauvinistically in favor of the English variant of this instinct as against the German. While McDougall (1908) likewise postulated a gregarious instinct, he did not derive all social conduct from this single source. The parental instinct, the instinct of appeal, and other instincts, too, have social reference. When about 1920 a sharp reaction set in against the instinct doctrine, gregariousness all but disappeared as a postulate from social psychology. Virtually no writer today employs it as an explanatory principle.

PRIMITIVE PASSIVE SYMPATHY

Before leaving instinctivist explanations, it is necessary to call attention to McDougall's theory of "the sympathetic induction of emotion" or, as he preferred to call it, "primitive passive sympathy" (1908). He states his view as follows:

> We must not say, as many authors have done, that sympathy is due to an instinct, but rather that sympathy is founded upon a special adaptation of the receptive side of each of the principal instinctive dispositions, an adaptation that renders each instinct capable of being excited on the perception of the bodily expressions of the excitement of the same instinct in other persons. (14th edition, p. 98)

In other words, if I perceive expressions of rage in another this perception is likely to arouse my own instinct of anger. There are two keys that will unlock each instinct: one is the "biologically adequate" cause; the other is the perception of the instinct in action in another person. The reader may rightly ask, Would not this formula lay all of us open to too much induced emotion? Should we not constantly be suffering instinctive seizures vicariously? McDougall escapes this dilemma in an interesting fashion. He brings laughter into the picture. Laughter is a device by which we avoid excessive sympathy. Byron wrote, "If I laugh at any mortal thing, 'tis that I may not weep." McDougall heartily agrees, and adds that laughter saves us not only from depression and grief, but from all other forms of vicarious sympathy as well.

Theodore Ribot. In his *Psychology of the Emotions* (1897) Ribot gives great prominence to sympathy, which he calls "the foundation of all social existence." He is not content with the two forms which most writers recognize, but postulates three. The first is the primitive or automatic type that we have suggested might be regarded as an instance of conditioned response. The second is reflective in that the individual is self-conscious concerning his state of mind. He knows that he feels for the other person even if the suffering is not his. The third is an intellectualized sentiment of loyalty, tolerance, or philanthropy — far broader in type than any specific instance of fellow-feeling. His three grades are similar to, if not identical with, Spencer's.

Ribot's approach is phenomenological. He is asking, What are the diverse forms of experience or "intentions" that men have, to which the single and inadequate label *sympathy* is applied? Are we dealing here with a simple unitary process, or with a term covering various types of mental acts?

Max Scheler. The phenomenological approach reaches its climax in Scheler's *Wesen und Formen der Sympathie* (1923). The German language is far more flexible than the English or French in dealing with this particular subject. A translation of this book was published by the Yale Press in 1954 under the title *The Nature of Sympathy.* There are, according to Scheler, eight forms of sympathetic orientation (or acts). For most of them clear English equivalents are lacking.

The first three are of a low grade, really "pseudosympathy," since in these cases the focus is upon the subject's own feeling and not upon his feeling *with* or *for* others.

1. *Einfühlung* is the primitive, reflex process mentioned by Smith, Spencer, Ribot, and others. The term *empathy* is a fair translation, provided it is understood to mean only elementary motor mimicry and is not employed in the broad sense of "a gift of understanding people" as is sometimes the case today.

2. *Miteinanderfühlung* calls attention to a condition of "simultaneous feeling," as in the case of two or more people reacting in a similar way to the same stimulus (e.g., an attentive and emotional cinema audience).

3. *Gefühlsansteckung* refers to the spread of feeling through social induction and facilitation, as in a mob, panic, or throng. It likewise includes McDougall's sympathetic induction of emotion. Such "transpathy" is still focused upon the affective experience of the individual, and is not truly sympathetic.

Higher levels of sympathy are found in:

4. *Einsfühlung*, wherein occurs an identification of feeling: the child playing with a doll identifies with her mother. Citizens in wartime feel united in suffering and in aim. There is an especially high degree of such "unipathy" between lovers who feel keenly the joys and sorrows of each other.

5. *Nachfühlung* is far more conscious and detached. We encounter it in the statement, "I know just how you feel." In such a case, however, we distinguish clearly our own feeling from another's. We may even add, "I know how you feel, but I wouldn't act as you are acting." This state of mind is Adam Smith's "intellectualized," or Ribot's "reflective" sympathy.

6. *Mitgefühl* may best be translated as "fellow-feeling." It is phenomenologically distinct from all states of self-reference, and denotes that clear and normal condition where one person appreciates accurately another's feelings, and usually (though not necessarily) values them positively. Fellow-feeling is the primary basis of social living — a product of phylogenetic evolution.

7. *Menschenliebe* is the level of sentiment recognized explicitly by Ribot and McDougall. One not only senses the other person's state of mind, but prizes it and respects it. Altruism and philanthropy belong to this class.

8. *Akosmistische Person- und Gottesliebe* (unworldly sympathy) refers to the sense of mystical sympathy that characterizes some people's religious orientation to life as a whole — "a unity of all finite spirits in God."

Scheler's phenomenological analysis has not been surpassed in subtlety. Besides being valuable in its own right, it illustrates the important fact that psychological analysis depends to a considerable degree upon the categories of language available to the analyst.

Empirical work. It was not until the decade of the 1930's that sustained attempts were made to reduce the problems of sympathy to an empirical basis. Freudian theory, World War I, and the prevailing temper of irrationalist theories had focused attention earlier upon aggressive behavior. But during the business depression interest in cooperative conduct rose. A summary of work prior to the late 1930's is found in May and Doob (1937), and pertinent anthropological evidence is presented by Mead (1937).

Employing the method of observation, Lois Murphy (1937) analyzed five thousand episodes of behavior in a nursery school. While aggressive acts were much more numerous than sympathetic acts, she found that the latter occurred with considerable frequency, chiefly among children who felt psychologically secure, who themselves had had an experience similar to that evoking their sympathy, and among children who came from certain types of cultural surroundings.

In recent years considerable relevant work has been added, although the concept of *sympathy* or even of *cooperation* is seldom used, probably because their flavor is moralistic rather than scientific. But much of the current research in group dynamics, in industrial relations, in the psychological aspects of group tensions, and in psychotherapy represents an empirically action-oriented extension of the movement we have here surveyed. The interested reader will find guidance in Leighton (1945), Pear (1950), Chase (1951), Maier (1952), Sorokin (1950).

IMITATION

Social psychologists are required to explain the overwhelming fact of social conformity in human behavior. Parents set models to which children conform; fashions are models to which adults conform; and culture itself is a model to which everyone (or nearly everyone) conforms. No problem in social psychology is more insistent. The term "imitation" designates the problem — but does not solve it.

Certain nineteenth-century writers, however, saw in imitation a simple and sovereign principle suited to explain all conformity. One single force was thought to provide both the motive and the means. Walter Bagehot (1873) represents this point of view:

> At first a sort of "chance predominance" made a model, and then invincible attraction, the necessity which rules all but the strongest men to imitate what is before their eyes, and to be what they are expected to be, moulded men by that model. (1875, p. 36)

Imitation is thus the conserving agency in society accounting for the "cake of custom" that

1268

binds all but the strongest men. It accounts for the "sameness" in savage societies, for the mimicry of children, for the gregariousness of species. Since conformity behavior is the result of "invincible attraction" the really critical problem lies in invention and novelty. Inventive departures from conformity Bagehot accounted for by the activity of "strong men," whom he also called "nation builders."

More extended is the treatment given by the French author Tarde, 1843-1904. For him, as for his British contemporary, imitation was "the key to the social mystery." Enthusiastically he asserts "Society is imitation" (1903, p. 74). Among Tarde's famous laws of imitation are the following: (a) The law of *descent* maintains that superior classes are imitated by the socially inferior. Fads, we might note today, start in Paris or on Park Avenue and die on the counters of the dime stores. (b) The law of *geometrical progression* calls attention to the rapid dissemination of a fashion, rumor, or craze from the point of origin. (c) The law of the *internal before the exotic* helps to account for the fact that one's own culture is imitated in preference to foreign cultures. About the process of imitation itself Tarde has little to say. He likens it to a hypnotic dream state — to somnambulism, which runs its course under the dominance of photographic images formed by the model in people's minds. This formulation runs parallel to the concept of "suggestion," equally prominent in French social psychology of the same period (Charcot, LeBon).

Novelty and invention, according to Tarde, are the result of conflicting imitations. All opposition is a matter of clash between contrary models. Competition, discussion, war are the three great forms of social opposition; they arise because contrary models are being imitated.

Unlike his countryman Durkheim, Tarde was an individualist. He once wrote to Baldwin in America, "your point of arrival is my point of departure," by which he meant that Baldwin's treatment of imitation as a specific mechanism in each individual person is the foundation upon which Tarde himself would build his more sweeping sociology of imitation.

Baldwin found in imitation the key to mental development in the child (1895, 1897). His account calls attention to the fact that there seem to be at least two forms of imitation — nondeliberate and deliberate (we might say today conditioned reflex and insightful). The process involves three stages. In the *projective* stage the child receives impressions of a model as a photographic plate receives an image. (The meaning of "projective" here differs from the modern use of the term.) In the *subjective* stage the child tends to assume the movements, strains, attitudes of the model. He is "a veritable copying machine" and cannot help doing so. It should be noted that at this crucial stage in the process Baldwin seems to rely on the little understood tendency to elementary motor mimicry. He also calls attention to the large amount of self-imitation in infancy — circular reflex activity — which makes for the repetitiousness of the child's behavior. Finally, in the *ejective* stage the child reaches a comprehension of the model; he may recognize that he is acting like the other; he knows how the other feels. Here imitation is an avenue to knowledge of the *alter*.

Mead (1934) saw clearly the deeper implications of Baldwin's ejective stage. It helps to explain how the individual comes to take the role of other people. We do more than imitate. We perceive what the other is doing (through ejective imitation) but we also perceive our own response to it. What occurs is an interweaving process. Each act is the resultant both of our role assumption and of our self-perceptions. There follows, therefore, a "conversation of gestures" from which emerge mutual understanding and a continual accommodation and adjustment of two people's sets of interests. In the process, language is of the highest importance, and "significant symbols" become a substitute for gross motor adjustment. Mead's theory of socialization is not exclusively "imitationist"; yet his point of departure lies in the doctrine as expounded by Baldwin.

Reference should be made to the tendency of many writers to include an "instinct of imitation" among the potpourri of motives postulated in man's original nature. William James did so (1890). Bernard (1924) shows that the acceptance of an instinct of imitation as one innate motive among many others, has been a common but sterile practice in the history of social and psychological science.

The writers we have mentioned, with the possible exception of Mead, regard imitation as having motivational force. Later writers tended to shift their attention away from the "why" of imitative behavior to the "how." Imitation is regarded merely as a means by which other motives reach their goal. McDougall perceived this issue clearly. He felt that, unlike true instincts, imitation had little or no urge character, and no specific goal. It seemed to him to be rather a "nonspecific innate tendency" (1908, Chapter 4). In this respect it should be classed with play, and also with "primitive passive sympathy" which we have previously discussed. Yet McDougall was much troubled by the process of precision which enables a child (or a parrot) to accommodate his vocal sounds to a model; as well as by the manifest tendency of spectators to assume the postural strains of the dancers or athletes they are watching; also by the child's tendency to take over the tensions of an adult into his own motor (and emotional) conduct. Blanton and Blanton (1927) emphasize the importance of postural tensions in the mother which tend to be assumed even by an infant in arms, so that a mother's fear becomes the child's fear.

This process of *empathy* remains a riddle in social psychology. It would seem to be genetically and conceptually basic to social learning and to lie at the heart of any theory of imitation. Some motor mimicry, as we have said, seems reducible to previous conditioning, but in other cases it appears to precede and to be a precondition of learning. The nature of the mechanism is not yet understood. It is obvious that the process of perception itself entails motor adjustments (e.g., of the eyes) that are "imitative" of certain properties of the perceived object. This fact may give the clue we need. The problem is discussed further by G. W. Allport (1937, pp. 530-533) and by Lipps (1907).

Most accounts of the "how" of imitation do not reach back to this simple empathic phenomenon. Many authors lean on the "ideo-motor theory," which Cooley states as follows:

> It is a doctrine now generally taught by psychologists that the idea of an action is itself a motive to that action and tends intrinsically to produce it unless something intervenes to prevent. This being the case, it would appear that we must always have some impulse to do what we see done, provided it is something we understand sufficiently to be able to form a definite idea of doing it. (1902, pp. 62 f.)

The ideo-motor theory is no longer "generally taught." But its potential value as an explanation of imitation (if it were true) is great. It merely says that all ideas press toward expression. Since many of our "ideas" come while we are observing other people (or things) it follows that we tend to act out (imitate) our perception of the stimulus pattern. While we cannot say that this theory is without merit, it is coarse and needs considerable refinement and modernization.

When idea-psychology declined and behaviorism arose an explanation of imitative behavior was sought in the properties of the conditioned reflex. Classical (Pavlovian) conditioning is represented in the theories of Humphrey (1921), F. H. Allport (1924), and Holt (1931). Let us take Holt's statement of the matter in his formulation of what he calls the "echo principle."

> A child will learn to echo back any action of another provided that another's performance of the act stimulates any of the child's sense organs at a moment when the child is engaged in a (random) performance of the same act. (1931, p. 112)

To illustrate, let us take the infant's early game of pat-a-cake. The theory assumes that first the child is moving his hands randomly in a clapping manner. The fact that he does so repetitively is an important initial fact, for the circular reflex involved provides a longer base in time for the conditioning to take hold. Perceiving this action, the parent likewise claps his hands (note that the parent is the first "imitator") and says "pat-a-cake." Thereafter the child tends to clap his hands when he sees the action, or hears the words "pat-a-cake."

While trying to keep the simplicity of a stimulus-response view of the matter, later authors regard the formula of classical conditioning as inadequate, and prefer to look at the phenomenon in terms of *instrumental* conditioning or "reinforcement." According to this approach, imitation is learned because it brings rewards and satisfactions. The young boy, for example, will develop a generalized tendency to imitate his brother when he finds that candy, comfort, and other desirable goals follows his (at first accidental) imitation of his brother's behavior. This "matched-dependent" relationship is regarded by Miller and Dollard (1941) as accounting for most forms of imitation. It encounters some difficulties when applied to the process of deliberate conscious copying, and does not seem to apply to the elementary phenomenon of motor mimicry in infants (prior to reward or satisfaction).

Gestalt psychologists (Köhler, 1927; Asch, 1952, Ch. 16) doubt that a conditioned imitative tendency can be established even with the aid of reward. Apes as well as human beings seem normally to require a *comprehension* of the means-end relationship represented by the model before they adopt it. An organism directed to a goal will reproduce in the motor sphere what he sees in the model *only* if he has familiarity with the situation and sees the relevance of the model for his goal. In this type of theory we are dealing with conscious and deliberate copying, which to some psychologists presumes "insight," or some other "cognitive" principle of learning.

Turning from the cognitive to the emotional level, we encounter the familiar concept "identification." The term obviously refers to *affective* imitation. The boy identifies with the father, the bobby soxer with an actress, the Negro with his race, and the humanitarian with mankind. This broad term is the key concept in psychoanalytic explanations of socialization (conformity) of the child. He is said to learn his morality and basic character-structure through emotional imitation of the parent and other authority figures. Freud (1921) makes the interesting suggestion that people who have no particular emotional significance for us are understood through empathy, but those who are of emotional value to us are understood through identification.

Finally, we call attention to the importance of imitation in cultural anthropology. Wissler has said that without assuming some imitative process it would be hard to imagine how men could possess a culture, "for it must be perpetuated by the imitation of the older by the younger members of the group" (1923, p. 206). The "diffusionists" in anthropology are accused of exaggerating the extent of imitative adoption of culture traits. Anthropologists who argue for "multiple origin" say that much apparent imitation is specious. They say that similarly constructed people with similar needs *would* make the same inventions and do the same things — without imitation being involved. The same caution can extend to everyday behavior. Children and adults may do the same thing under similar provocation without in reality imitating at all.

Summary. The reader who finds this survey uncomfortably condensed will profit from reading the second Appendix to Miller and Dollard, *Social Learning and Imitation* (1941).

Imitation is a protean concept. Its definition must inevitably remain exceedingly broad. *The term refers to any occasion where a stimulus situation gives rise to motor activity of a sort that resembles the stimulus situation.* Since various types of processes may have this broad result, we may well suspect that there is no single entity. Also we may suspect that no one "learning theory" will account for the various forms of imitation that grow up in human conduct. Investigators and writers have called attention to at least five apparently distinct mechanisms that may be involved:

(1) *Motor mimicry, empathy* — basically a perceptual motor reaction at present not fully understood (Lipps, McDougall, Blanton and Blanton).
(2) *Classical conditioning,* including the echo principle (Holt, Humphrey, F. H. Allport).
(3) *Instrumental conditioning,* emphasizing the rewards involved in engendering either specific or generalized imitative habits (Miller and Dollard).
(4) *Cognitive structuring,* including all instances of deliberate copying, and insightful reproductions (Köhler, Asch).
(5) *Identification* — having at present various meanings, but emphasizing the emotional disposition of the whole person to resemble some model (Freud).

While it is not impossible that these processes may in time be successfully reduced to a single formula, the prospect seems most unlikely.

As for earlier writers on the subject, we may declare that their interest lay primarily in the societal consequences of imitation. Though they tended to postulate a sovereign motive and mechanism, their real concern was with the social results of conformity (Bagehot, Tarde, and, to some extent, Baldwin, Mead, Cooley). Eventually, of course, the sociology of conformity will receive illumination from the mechanisms of imitation when these are more fully understood, but for the time being it is well to view the problems as distinct. The sociology of conformity must wait for an improved psychology of the processes that result in similar or same behavior.

SUGGESTION

As we have pointed out, sympathy, imitation, and suggestion are the Big Three. Between them these concepts have dominated much of the systematic social psychology of the past. The differences among them, as we have said, reflect the preference of authors for emphasizing the functions of affection, conation, or cognition, respectively.

The cognitive flavor of the term "suggestion" stands out clearly in McDougall's definition:

> Suggestion is a process of communication resulting in the acceptance with conviction of the communicated proposition in the absence of logically adequate grounds for its acceptance. (1908, p. 100)

McDougall held that the dynamics of the process lay in men's instinct of *submission;* this instinct is aroused by any person or symbol having prestige, and accounts for our yielding to communicated propositions "in the absence of logically adequate grounds."

Historically considered, suggestion is the most important of the three concepts. This is true even though its impact on social psychology was not felt until about 1890. But so persuasive at that time was the work of Charcot, LeBon, Sighele, Sidis, and others, that nearly all the problems in social psychology became formulated both for scientists and for laymen in terms of suggestion. Even today the layman inclines to think of the whole field of social behavior as a matter of "mass mind," "mob hysteria," "power of suggestion." The concept brought about a marriage between abnormal and social psychology. While the union has cooled, it has not entirely dissolved.

The process of suggestion has been explained in various ways, chiefly in terms of:

animal magnetism (mesmerism)
ideo-motor response
dissociation of consciousness
association of ideas, redintegration, conditioning
reduction of determining tendencies
identification
cognitive restructuring

These explanations pertain to both suggestion and hypnotism, for most writers regard hypnotism as nothing more than an extreme instance of the suggestive process. The early history of these concepts is well reviewed by Boring (1929).

Animal magnetism was the vague formula offered by Mesmer (1779). By "magnetizing" a wand or a tub Mesmer found that when subjects in his seances touched these instruments they fell into trances and became obedient to his voice. Even in his day Mesmer's magnetic explanation was severely questioned. Yet his subjects did go into trances, and cures of neurotic ills did result. Mesmerism was a fact, however queer its trappings and theory.

Ideo-motor response. Braid (1843) rejected Mesmer, and invented the term "hypnotism" to describe the phenomenon under dispute. He noted the importance of sensory fixation. Hypnotism must start with a limitation of the field of consciousness. Sensory fixation is followed by a restriction of ideas, a process he called *monoideism.* An idea firmly planted in the patient's mind takes root and issues into motor behavior. Surgeons of the day were greatly interested, for in the era prior to anaesthetics it seemed a hopeful possibility that ideas planted in a patient's mind through hypnosis might dull the pain of the knife.

It is important to recall here that much nineteenth-century psychology was dominated by the concept of "idea." One heard of the mechanics of ideas, of the association of ideas, and of ideomotor action. Herbart, Kant's successor at Königsberg, was a dominating figure. His influence on educational practice was immense: nineteenth-century educational psychology became largely a matter of planting the right ideas (e.g., via copybooks) in children's minds, hoping that such ideas would issue into right actions. In this period Michael Faraday (1853) tried to improve "the morbid condition of public thought" by pointing out that table tipping and other spiritualistic phenomena could easily be explained by the subtle effects of the idea upon physical movement. The direct issuing of ideas of movement into action became known as *dynamogenesis,* and was regarded as a special case of the ideo-motor theory. The high repute of the theory as late as the time of William James is seen in his enthusiastic endorsement:

> We may then lay it down for certain that every representation [idea] of a movement awakens in some degree the actual movement which is its object. (1890, II, p. 526)

Liébeault (1866) published an important work in which he explicitly introduced the concept of *suggestion,* insisting that the process was identical with hypnotism, and that all people could under favorable circumstances be hypnotized. To him suggestion was merely the acting out of an idea planted with rapport in a field of limited attention. This emphasis on *rapport* anticipates McDougall's claim that the submissive instinct is aroused in all cases of suggestibility, and also Freud's still heavier stress on identification as essential to the state of rapport that makes suggestible behavior possible.

Bernheim (1884) followed Liébeault in his thinking, and througn his direction of the Nancy school of psychiatry became influential in establishing suggestive psychotherapy based on the view that suggestion is an essentially normal ideo-motor process.

Dissociation of consciousness. Opposition to the theories of the Nancy school sprang up in Paris at the school of Salpêtrière, headed by Jean Martin Charcot (*Oeuvres complètes,* 1888-1894). As a persuasive and dramatic clinician, Charcot demonstrated to gaping audiences the phenomena of the hysteric splitting of personality under hypnosis. Unlike the Nancy school, he maintained that only hysteric personalities could be hypnotized (hence hypnosis and normal suggestion do not form a continuum). Some of Charcot's students, e.g., Pierre Janet and Boris Sidis, modified this extreme position. Sidis (1898), for example, regarded hypnotism as "abnormal suggestibility," marked by deep dissociation and disaggregation of consciousness. By contrast "normal suggestibility" falls short of true hypnotism, but is also a matter of splitting off (to a slighter extent) the waking, guiding, controlling, guardian consciousness from the automatic, reflex, subconscious self. Abnormal suggestibility can be induced through direct commands ("You are falling asleep"), whereas the normal person tends to resist direct command and to respond only to subtle indirect suggestions introduced at "the psychological moment" ("Tommy, it's bedtime"). Sidis sums up his views under three laws:

1. Suggestibility varies as the amount of dissociation and inversely as the unification of consciousness.
2. Normal suggestibility varies with indirect suggestion and inversely with direct suggestion.
3. Abnormal suggestibility varies with direct and inversely with indirect suggestion.

Charcot's influence on clinical psychiatry was, of course, very great, both through students who followed his teaching closely (e.g., Morton Prince) and those who departed from it, or reacted against it (Janet, Sidis, Binet, Freud). Equally great was his influence upon social psychologists, who found in the doctrine of dissociation an explanation for crowd phenomena, mobs, mass hysteria, demagogic leadership. Perhaps the most influential book ever written in social psychology is LeBon's *The Crowd* (1895). It is a direct product of Charcot's teaching. Later we shall speak of it more fully.

This school of thought took a grim view of man's subconscious nature. Whether released through hypnosis, hysteric seizure, or mob conditions, the automatic, subconscious self is an antisocial actor. In 1886 Robert Louis Stevenson immortalized the Charcotian view in *The Strange Case of Dr. Jekyll and Mr. Hyde,* clear evidence that the dramatic teachings of the Salpêtrière school had by that time become international public property. A decade or two later, in spite of Freud's rejection of Charcot's formulations, his concept of the unconscious Id took over the same essential function as Charcot's split-off subconscious self. This concept likewise became public property, and has helped maintain the supremacy of irrationalism in clinical, academic, and popular psychology to the present day.

Starting about 1890, many writers, as we have said, enthusiastically applied the concept of dissociation to social behavior, thus bringing about the marriage between abnormal and social psychology. Charcot's pupil, Morton Prince, found it entirely natural to enlarge the scope of the *Journal of Abnormal Psychology* which he had established in 1908, and to rechristen it in 1922 as the *Journal of Abnormal and Social Psychology*. He regarded the two fields as indissolubly linked.

Association, redintegration, conditioning. A more conventional type of theorizing keeps suggestion safely within the confines of simple associationism. A stimulus "suggests," i.e., touches off, a particular idea (determining tendency, brain path) that leads to a particular response. For a fisherman a sunny day in springtime "suggests" fishing. The Scottish psychologist, Thomas Brown (1820), employed the term "suggestion" as wholly equivalent to "association." Bernheim (1884) of the Nancy school, while recognizing the importance of rapport and of ideo-motor response, also agreed that every impression, every mental picture, every association was a suggestion. Only slightly different is Titchener's statement that, "A suggestion is any stimulus, external or internal, accompanied or unaccompanied by consciousness, which touches off a determining tendency" (1916, p. 450).

A variant of this view makes suggestion a matter of *redintegration.* A previously formed disposition of a highly complex order can be rearoused as a whole by an associated cue. The sunny springtime day is the cue: "going fishing" is the complex activated response. Hamilton (1859) first called attention to this part-arousing-the-whole relationship in association. More recently Hollingworth (1920) used it extensively to account for suggestibility in both normal and abnormal subjects.

The concept of the *conditioned response* is closely related to this line of thought. Suggestions are conditioned cues, and the process involved is simply the reactivation of a motor pathway to which the cue has been conditioned. "Tom, go to bed" (direct suggestion) or "Tom, it is bedtime" (indirect suggestion) are cues intended to reinstate Tom's nocturnal habit systems. This view is found in all behavioristic treatments of social phenomena, such as those of F. H. Allport (1924), Bechterew (1932), Hull (1934).

The manifest difficulty with these treatments is that they fail to distinguish suggestion from any other process of association of ideas, or from any other stimulus-response sequence. While it undoubtedly is true that suggestion should properly be viewed as a subclass of associative phenomena, still to have distinctive value as a concept there must be some *differentia specifica* to give it its peculiar complexion among such phenomena.

Reduction of determining tendencies. The differentiation required is stated in Warren's Dictionary (1934):

> A suggestion is a stimulus, usually verbal in nature, by which one individual seeks to arouse activity in another by circumventing the critical, integrative functions.

We have already cited McDougall's similar definition, which refers to the acceptance of a proposition without logically adequate grounds. In these and in similar definitions, attention is called to some arbitrary restriction in the determinants of behavior. The individual is not employing all relevant ideas, nor his full intelligence. He is for the moment acting without full self-determination. Granted that suggestion proceeds according to the laws of association (conditioning), still we must allow also for the *blocking* of normal associations, so that the end result in behavior is due to a selected field of determinants. Braid, we have noted, labeled this restriction "monoideism." When we speak of "reduction of determining tendencies" we are not, of course, referring to the sheer numbers of associations present, but to a functional restriction of the brain field, so that the "critical, integrative" functions are circumvented.

Let us take the case of a man who buys a suit of clothes under the persuasion of a silver-tongued salesman. The seller points out its virtues in terms of fashion, flattering fit, and comfort. Bombarded by these arguments the customer yields to the persuasion. But the buyer has not given due weight to equally relevant considerations that might have blocked the purchase: expense, specific wardrobe needs, and personal preference. Even though the salesman may have introduced *more* determining tendencies than the buyer initially possessed, yet functionally speaking, the salesman has reduced the *effective* determinants. The buyer's purchase is, therefore, partly, if not altogether, the result of suggestion.

Around the turn of the century a great deal of psychological thought and writing was ordered to the concept of suggestion. Many books and articles dealt with automatisms, dissociation, hysteria, and like matters. Against this background Binet (1900) made two significant contributions. He reacted against hypnosis, agreeing with Wundt that its use was immoral, and sought instead for "inoffensive" methods by which to explore "normal suggestibility." In so doing he leaned away from Charcot and toward Bernheim. His second contribution had to do with his invention of experimental methods. He presented children with lines, with drawings, with questionnaires in such a way that some strong "idées directrices" were planted in their minds. Later they were asked to reproduce the stimulus, with the result that specific distortions occurred in the direction of the implanted idea. In group behavior he observed the effects of prestige among children. The leaders in his experimental groups of children proved to be less suggestible than the followers. Owing largely to Binet's efforts, suggestion became a standard problem for academic psychology as well as for clinical.

One favorite problem had to do with the conditions of suggestibility. Who is suggestible? The answers given, as summarized by Ross (1908) include several dubious assertions. Certain species of animals (e.g., sheep) are more suggestible than others; certain "races" (Frenchmen and Slavs) are more suggestible than others (Anglo-Saxons); children more so than adults; women more so than men; people are suggestible if they are temperamentally absent-minded or in a state of emotion, such as jealousy; they are vulnerable toward sources possessing prestige; and are especially suggestible in crowds, or toward often repeated assertions.

There is something rigid about this approach. Krech and Crutchfield (1948, p. 333), for example, argue that suggestibility is not a general trait. No woman is uniformly more suggestible than a man. Indeed, in areas where she is well informed (e.g., in shopping or child rearing) she is probably less suggestible than her husband.

Cantril has clarified the matter. He states that suggestion occurs especially when the individual is confronted by a "critical situation" where he cannot readily make a decision. He will accept the suggested proposition for belief (and action):

> (1) if he has no adequate mental context of his own for the interpretation of the event; or (2) when his mental context is so rigidly fixed that the suggestion automatically arouses this context (e.g., a prejudice) and the person fails to examine the situation in its own right. The first condition results from bewilderment; the second from the "will to believe." (1941, pp. 64 f.)

To paraphrase this view, people will act uncritically (with a reduction of determining tendencies) if they are relatively unfamiliar with a topic, unaccustomed or unable to check up on the suggestion offered to them — in short, if their mental organization is *unstructured.* They are likewise prone to accept the suggestion if it fits into their rigidly structured stereotypes, thus nourishing and calling into action pre-existing attitudes, prejudices, beliefs. An "unstuck" mind and an "overstuck" mind both favor suggestibility.

It would not be hard to cast the entire psychology of crowd control, demagoguery, and propaganda into this framework of thought. The devices used are often of a type to "unstructure" people's minds. Demagogues call attention to the bad state of affairs, to the malfeasance of government officials, or to the threats of poverty, ill health, unpopularity. When people are sufficiently unsettled by this barrage, the propagandist states his remedy. It is then likely to be accepted as a way out of the confusion. Or, still more frequently, demagogues appeal directly to people's "overstuck" convictions. The propagandist, orator, or advertiser simply associates his product (or message) with pre-existing idols (sentiments, prejudices, convictions) and thus has it accepted because it is congruent with the person's pre-existing desires and aversions. George Bernard Shaw once defined propaganda as "the organization of idolatory" (the organization of the propagandist's message with an old idol).

Identification. McDougall, as we have seen, invoked the submissive instinct as a necessary link in the suggestion process. The Nancy school felt that good rapport was necessary between *suggestionneur* and *suggestionné*. Freud felt both these views to be pallid. Thoroughly familiar with the theories and therapy of Nancy and Salpêtrière, he rejected completely their concepts of suggestion and hypnosis. There are deeper reasons, he said, why a leader attracts followers and is able to mold their behavior. It is because they find in him a love object, one who relates himself to their thwarted sex needs (1921). In the case of male groups and crowds (e.g., Nazi storm troopers) a deeply repressed homosexual interest is revealed through their passivity toward, and identification with, the leader. What has been called suggestion, especially in mass situations, is in reality a consequence of identification. The theory has been more fully explained, with special reference to the leader-follower relation in wartime, by Waelder (1939). Its merit lies in calling attention to the importance of emotional attachments in bringing about obedient behavior. It helps also to account for the fact that members of groups, through a common focus of identification, also identify with one another. It used to be said in Germany that there is no such thing as a "single Nazi." Only with the support of a group does the peculiar subservience to the leader and his ideology take possession of the individual.

As cognitive restructuring. In recent years considerable skepticism has been expressed concerning the utility, and even the validity, of the concept of suggestion. Asch, for example, has declared flatly that "the case for the doctrine of suggestion has not been proven" (1948, p. 251).

The argument turns on the question whether there is any real difference between conduct called suggestible and conduct that is said to show critical ability or full self-determination. Another way to state the issue is in terms of "rationality." Is the person who seems to us to be suggestible, in reality behaving any less rationally than the person who seems to us to be nonsuggestible? From their own points of view, both are probably acting in an intelligent way, making the best use they can of the totality of information that confronts them.

One experimental example will suffice to etch this new point of view (which derives from phenomenology and Gestalt psychology). Subjects are asked to think about the following passage, and are told that its author is Thomas Jefferson:

> I hold it that a little rebellion, now and then, is a good thing, and as necessary in the political world as storms are in the physical.

When asked whether they agree with the sentiment expressed, and what it really means to them, subjects generally approve it, and interpret the word "rebellion" to mean somewhat minor agitation. But when subjects are told that the author is Lenin, they ordinarily repudiate the statement, and interpret the word "rebellion" to mean violent revolution (Asch 1952, pp. 419-425).

Just what is happening here? The doctrine of suggestion would say that people hold previous emotional attitudes toward Jefferson and Lenin, the former being "good," the latter "bad." Hence the statement when ascribed to Lenin is rejected, for it is not viewed critically in its own right, but is responded to only in terms of one's prejudice against Lenin. The selfsame statement is accepted when it "redintegrates" one's veneration for Jefferson. Thus the person is suggestible, reacting not to the statement in its own right, but on the basis of preformed attitudes toward its presumed author.

The cognitive theorist, however, points out that the author's name altered the whole *meaning* of the stimulus. This fact is shown by the subject's tendency to interpret "rebellion" in different ways. If Jefferson wrote the passage he could only have meant by it mild agitation; if Lenin, he surely had in mind violent revolution. Hence it is *not* the "selfsame" passage, evaluated differently because of the suggestive force of the authors' names. Rather, two dissimilar stimuli are evaluated rationally (i.e., according to the true meaning of the stimulus in context). As Asch puts it, it is not the judgment of the object that is altered (through suggestion), but the object of judgment is itself altered, and is interpreted objectively and rationally.

Whether this line of argument can be extended so far as to eliminate the need for the distinctive concept of suggestion is doubtful. Even in the case cited it seems clear that people's prejudices for Jefferson or against Lenin are active factors in the process of interpretation, and that there is a tendency therefore to react to the statement not with all one's critical associations, but primarily on the basis of the pattern of reduced determining tendencies activated by the author's name.

The term "cognitive restructuring" does indeed describe (phenomenologically) what seems to happen in all cases of suggestion. We perceive and interpret matters differently in different circumstances. But whether "cognitive restructuring" is an *explanatory* concept is a question. It serves to warn us that much behavior that seems suggestible and irrational to others is not so from the actor's point of view. He is doing what he can in the given circumstances. Yet the fact remains that in *certain* circumstances his behavior is far more critical and self-determined than in *other* circumstances. Hence, we still need the concept of suggestion to designate the distinction in behavior that results from a narrow, rather than a full, set of determinants.

Summary. Most, if not all, of the views here surveyed contain some truth. We start with the simple observation that in certain circumstances people act with less critical capacity than in others. When their critical sense is wholly or partly in abeyance we are likely to find a high degree of dominance of one idea or image (monoideism). Dynamogenic tendencies may also be present. Redintegration and conditioning help touch off the decisive determiners of conduct. In extreme cases there may be a hysteric splitting of the personality. In all cases a cognitive restructuring takes place, in the sense that the individual organizes the determinants that are acting upon him. This organization can often be controlled from outside if the proper conditions are present (e.g., an unstuck or overstuck mind), and if leaders or manipulators use adroit appeals, aided perhaps by bright trappings (sensory fixation), and eloquence. Rapport, identification, a submissive attitude, and habits of obedience greatly facilitate the process.

Thus suggestion is a complex phenomenon. We need all the insights that history has bequeathed us to construct an adequately comprehensive theory.

CROWD

Plato feared democracy because he feared the irrationality of men. In relative solitude they may deliberate wisely, but taken herdwise their reasoning is defective. "Had every Athenian citizen been a Socrates, every Athenian assembly would still have been a mob." While this suspicion of human collectivities is ancient and recurrent, it was not until the 1890's that a systematic theory of crowd behavior became available.

Le Bon's reliance on suggestion, considered as a hysteric splitting of personality, is seen in the following quotation. In a crowd, he writes,

> the individual may be brought into such a condition that, having entirely lost his conscious personality, he obeys all the suggestions of the operator who has deprived him of it, and commits acts in utter contradiction with his character and habits. (1896, p. 34)

Crowd conditions release deep prejudices, racial tradition, and brute instinct. The crowd man

> is no longer conscious of his acts. In his case, as in the case of the hypnotized subject, at the same time that certain faculties are destroyed, others may be brought to a high degree of exaltation. Under the influence of a suggestion, he will undertake the accomplishment of certain acts with irresistible impetuosity. (p. 35)

People in crowds and crowd leaders are given to *action*, never to critical *thought*. When a man joins a crowd he "descends several rungs in the ladder of civilization. Isolated, he may be a cultivated individual; in a crowd, he is a barbarian, that is, a creature acting by instinct" (p. 36). Since Le Bon believed in the excessive suggestibility of women and of certain races, he adds, "Crowds are everywhere distinguished by feminine characteristics, but Latin crowds are the most feminine of all" (p. 44).

Having applied the theory of dissociation to crowd members, Le Bon next considers the attributes of the crowd leader. He is a person who knows how to invoke pointed and clear images "freed from all accessory explanation." Crowds act only on the basis of "image-like ideas." While these may be accidental and passing — such as an auctioneer or barker might employ to make a sale — more often they are fundamental, deep-lying convictions of an instinctive or quasi-religious order. But in any event the leader will deal only in image-like ideas that point directly to action (ideo-motor theory). The leader's immediate tool in the control of ideas is the well-chosen word. "The word is merely the button of an electric bell that calls them up" (p. 118). The word evokes the image, the image evokes a sentiment, the sentiment leads to action. Crowds "are not influenced by reasoning, and can only comprehend rough-and-ready associations of ideas. . . . The laws of logic have no action on crowds" (p. 128). The clever leader will obey the principles of affirmation and repetition. He will assert, not argue; and above all he will, as Hitler later also advised, give his proposition thousandfold repetition. In discussing the elementary principles of crowd control Le Bon set the stage for modern work on demagogic leadership (cf. Lowenthal and Guterman, 1949) and on the rules of propaganda (cf. Doob, 1935).

It is obvious that Le Bon had a low opinion of the crowd-man. He saw crowds as a menace to established institutions. Crowds flourish as civilization declines. He wrote "The age we are about to enter will in truth be the era of crowds." Masses are bent on destroying society as we know it and harking back to "primitive communism." Was Le Bon thinking prophetically of the fascist and communist upsurges of the twentieth century? Yet for all his apprehension and conservatism Le Bon was unwilling to condemn democracy. While he deplores the irrationality of electoral crowds, of juries, of parliaments, he thinks it well to risk their errors, for in their way crowds do express the basic aspirations of the race, and at times act with a rough heroism.

Sighele's evaluation differed. This Italian writer — whose basic theories were so much like Le Bon's that they fell into a great dispute over priority — said that democracy, because of its crowdish tendency, was an evil. As an Appendix to his *Psychologies des Sectes* (1895) he published an essay *Contre le Parliamentairisme* which laid one of the cornerstones for later fascistic ideology. Parliaments were crowdish — away with them! Curiously enough Sidis (1898), an American, drew opposite conclusions from similar premises. Like Le Bon and Sighele, he held suggestion to be a basic evil, but he felt that democracy could and should combat its ravages through proper education, so that eventually the decisions of the electorate and of parliaments might be reached on a more rational basis.

There is a systematic character in Sighele's work that merits attention. His three volumes form a logical series: *Le crime à deux* (1893), *La foule criminelle* (1891), and *Psychologie des sectes* (1898). All are saturated with the evils of suggestion. Man's subconscious nature is bestial and criminal. Hence suggestion and criminality are inevitably linked. Like Lombroso and many of his countrymen, Sighele was captivated by the study of crime and other abnormal manifestations of social life. But what is noteworthy about this author is his orderly conception of social psychology. He deals with the dyadic relationship — two individuals bound together as *suggestionneur* and *suggestionné,* also with the larger transitory collective of the crowd, and eventually with more organized and enduring groups or *sectes.* No other social psychologist seems to have focused his work at the basic level of the dyad, and then tested his views in relation to progressively larger groupings. The idea is a good one. The dyadic relationship has, however, received extensive treatment by Simmel (1950).

Both Le Bon and Sighele maintain that crowd phenomena do not necessarily require physical proximity. While most crowds are *congregate,* they may (especially in these days of radio communication) be *consociate.* Even in his day Le Bon observed that

> Thousands of isolated individuals may acquire at certain moments, and under the influence of certain violent emotions — such, for example, as a great national event — the characteristics of a psychological crowd. (1896, p. 27)

While it is certainly true that mass phenomena are ordinarily found only in congregate crowds, we must allow for exceptions. There are on record at least two cases of a consociate panic — the result of Orson Welles' broadcast *The War of the Worlds* in the United States in October, 1938 (Cantril, 1940), and in Equador in February, 1949 (Britt, 1950, Section 31).

The combined influence of Charcot, Le Bon, Sidis, Sighele, Tarde, blew like a gale upon the American sociologist, E. A. Ross. His *Social Psychology* (1908) brought to a climax this whole vigorous movement. While Ross deals with many of the consequences of suggestion (fashions, crazes, conformity, custom), he regards crowds as the crowning mass phenomenon. In them the strength of "multiplied suggestion" is at its maximum. The individual is "helpless to control his position or movements." Ross's account of crowds is strictly in the tradition we have described, and shares its atmosphere of melodramatic exaggeration.

Turning our attention for a moment from congregate crowds to the cognate phenomenon of "crazes," let us list briefly Ross's laws (1908, p. 76):

1. A craze takes time to develop to its height.
2. The more extensive its ravages, the stronger the type of intellect that will fall prey.
3. The greater its height the more absurd the propositions that will be believed.
4. One craze is frequently followed by another.
5. Dynamic society (e.g., American) is more craze-ridden than one in the ruts of custom.
6. The higher the craze the sharper the reaction against it.
7. Ethnic or mental homogeneity is favorable to the craze.

Let the reader ask himself the following questions. What would be a good example to illustrate each of these various laws? On what basis did Ross devise them? Are they well-framed scientific propositions? Why, or why not? If the reader will take the time and trouble to carry through this exercise, he will then be in a position to evaluate properly the influential suggestion-imitation school of thought that we have been discussing.

E. D. Martin's *The Behavior of Crowds* (1920) leans on both the Le Bon-Ross tradition and on Freud. From the former the author derives the conviction that "the crowd mind is a phenomenon which should best be classed with dreams, delusions, and the various forms of automatic behavior" (p. 19). From the latter he derives his belief that crowds are agents for the release of repressed impulses. There are hunger crowds, fear crowds (panics), aggressive crowds (mobs), and other emotional masses, each one corresponding to some one long-standing repression that is struggling for an outlet. Under favorable conditions crowds offer a channel for their release. In this sense, therefore, crowds are a safe way for men to "go crazy together." While Freud (1921) would agree in part, we have seen that his theory of suggestion and crowd requires the added attribute of "identification" between the members and their leader.

THE GROUP MIND

Many authors, past and present, would agree that the interaction of individual minds produces a common manner of thinking, feeling, and willing, different from that of single minds in isolation, or from the mere summation of minds.

Yet the phrase "a common manner of thinking, feeling, and willing" is susceptible of various interpretations. In the period 1850-1930, roughly delimited, these interpretations were regarded as so many attempts to solve the problem of the "group mind." Solutions were offered by sociologists, philosophers, and by anthropologists, as well as by psychologists. In recent years the label "group mind" has fallen into disuse, but the same issues remain in new guises. For instance, what are today called "social institutions" are manifestly common manners of thinking, feeling, willing; so too is a "culture," also a "nation." The problem of the group and the individual, of the one and the many, is still with us. And the solutions of the past, often only slightly disguised, may be found in current theorizing.

A modern example. A team of five contemporary authors set out to discover "the functional prerequisites of a society" (Aberle, Cohen, Davis, Levy, Sutton, 1950). They defined society as

> a group of human beings sharing a self-sufficient system of action which is capable of existing longer than the life span of an individual, the group being recruited at least in part by the sexual reproduction of the members. (p. 101)

A nation or tribe is an excellent example of a society thus defined. Such a society, the authors argue, cannot endure unless the following conditions are met:

1. There must be provision for adequate relationship to the environment and for sexual recruitment of new members.
2. There must be role differentiation and role assignment (i.e., a division of labor to carry out the intricacies of the social organization).
3. Means of communication must exist.
4. The members must share a body of "cognitive orientations" (i.e., have common modes of thinking).
5. They must share certain goals (common modes of feeling and willing).
6. There must be a normative regulation of means (e.g., laws, ceremonials, customs).
7. Affective states must be regulated (because "the ungoverned expression of lust and rage leads to the disruption of relationships and ultimately to the war of all against all").

8. Socialization (i.e., new members must learn the accepted modes of behavior).
9. There must be effective control of disruptive forces (crime, fraud, promiscuity).

These authors write in an ahistorical vein; they make no mention of the venerable problem of the group mind. Yet the issues they treat are precisely those that have led many of their predecessors to argue for the existence of a group mind.

A close parallel may be seen between this list of functional prerequisites for a society and the criteria established for the group mind by McDougall (1920). (Numbers in parentheses call attention to the items in the above list that parallel McDougall's criteria.)

A. Continuity of existence, with the individual replaceable (1).
B. In the minds of members there must be an idea of the group (4,5).
C. Interaction with other groups, especially rivalry and conflict, promotes a group self-sentiment (implied in 3,4,5).
D. Existence of a body of traditions and customs in minds of members (implied in 3,4,5,6,7,8,9).
E. Specialization of functions and division of labor (2).

Whether we have made an exact paralleling for these two sets of criteria is not important. Our point is simply that historical formulations of the group-mind problem are still relevant to current theorizing. Earlier writers struggled to place the individual and his group in proper perspective — so do modern authors. Earlier writers sensed a continuity in the ideas, motives, habits of the group that was not dependent upon the ideas, motives, and habits of any particular individual. So do modern authors.

The social psychologist is not particularly concerned with the vast amount of ethnological and sociological evidence that may be adduced in support of the social mind. The existence of language, religion, tradition, nationalism is enough to convince him that a problem exists. What he wants to know is just how the social mind is supposed to operate. "Mind" is his province, and if in any sense at all there is a social or group mind he wants to know its nature and functioning. What then are the leading historical views from which he may make his selection — or upon which he must improve?

While any classification of theories is somewhat arbitrary, it seems helpful to distinguish seven types of emphasis. Occasionally a single author will be found to embrace more than one of these conceptions.

1. Analogical doctrines
2. Collective unconscious
3. Objective mind
4. Folk mind
5. Collective representations
6. Cultural determinism
7. Common-segment doctrines

Analogical doctrines. Throughout the ages men have likened society to an organism. We have already called attention to Plato's psychosocial analogy:

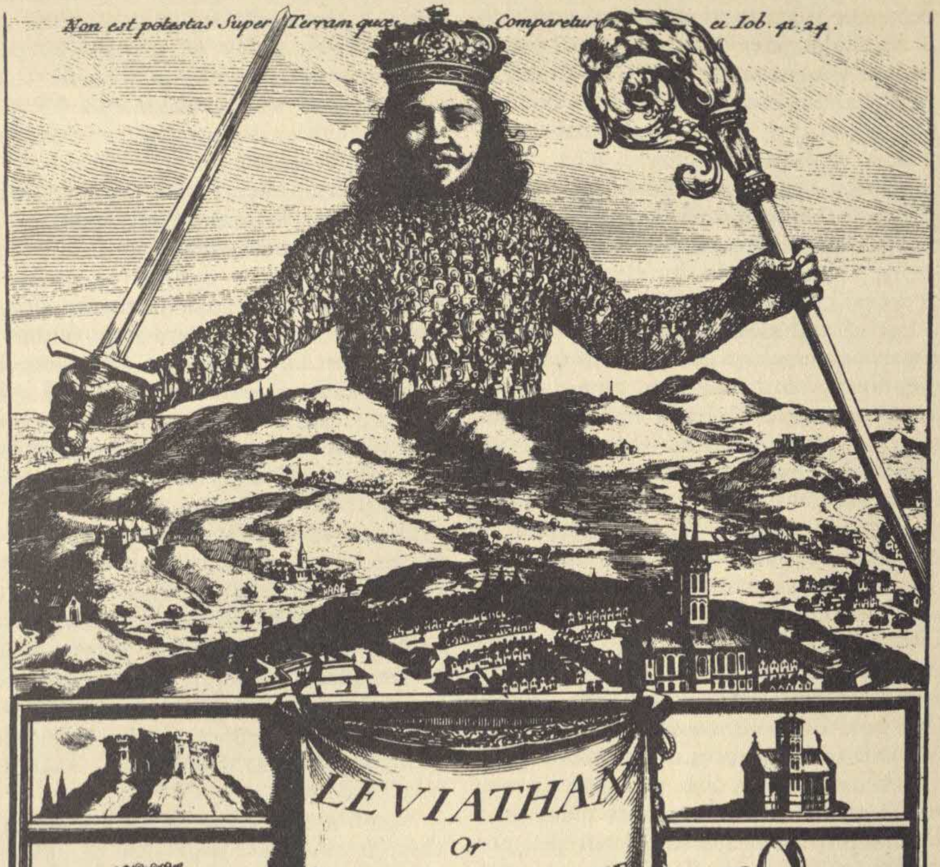
the head — cognition — the ruling class
the breast — conation — the warrior class
the abdomen — affection — the slave class

The first edition of Hobbes' *Leviathan* (1651) carried as its frontispiece a picture of a gigantic human figure made up of smaller men (Fig. 2). This simple imagery serves to arouse in our minds the problem that has troubled writers from Plato to the present day. How are we little men related to the Great Being of society? The first relationship that strikes us is one of *similarity.* Little men have thoughts, feelings, wills; so too in a sense does the Great Being.

Herbert Spencer (1876) pressed the analogy much further. The nutritive system of society includes agriculture and the productive industries; the circulatory system includes traffic, communication, the stock exchange; the regulating "brain" consists of the organs of government. With what we may suspect to be wry humor, Spencer ascribed to the House of Commons the functions of the cortex, to the House of Lords the functions of the medulla oblongata. Such analogies were vastly popular among nineteenth century social theorists.

Some writers, Spencer among them, pointed to limitations in the analogy: society lacks the specific external form that every organism possesses; its parts or "organs" are not in contact with one another and show greater freedom of mobility. Yet more convincing, these writers felt, are the similarities: the social organism, like the individual, has a continuous life history from birth to death; an increasing complexity marks its growth, along with an increasing specialization and differentiation of parts. What is most important of all, one cell (person) may die without disrupting the continuity of the whole. Individuals are replaceable; one perishes and one is born, but the common mental content and functioning of the group continue.

Analogical thinking is by no means a thing of the past. The organismic model still attracts many types of writers. One often reads such statements as the following: "Religion is the



M.2.T.
x

FIG. 2. "By Art is created the great Leviathan called a Commonwealth or State (in Latin Civitas) which is but an artificial Man, though of greater stature and strength than the Natural, for whose protection and defense it was intended." (From the Frontispiece to the First Edition of Hobbes' *Leviathan*, 1651).

heart in the organism of the group mind. If this heart ceases to beat, decay and decomposition set in." Mere metaphor — but helpful in making a point. A more subtle example of current analogical doctrine is seen in Wiener's conception of *Cybernetics* (1948). The initial model employed is modern mechanics. The "feedback" machines (e.g., thermostat) have servomechanisms like those of the human body. Both may, for example, become "jammed" through overloading or through faulty clearing. Similarly in society, systems of communication can be overloaded and clog, resulting in "noise." And just as neuroses and compulsions may obtain a monopoly on the nervous system of the individual, so may selfish tycoons buy up and dominate the communication channels of society. Enthusiasts for cybernetics might say that Wiener is really stating universal laws and thus forging a "unity of science"; the same laws of servomechanisms, communication, and noise apply to machines, to man, and to society. But others will see in cybernetics nothing much more than an artful revival of Spencerian thinking.

Collective unconsciousness. In a sense, all psychologists hold that our conscious mental life is continuous with a subconscious mental life. A few (not many) assert that this subconscious mental life reaches out to embrace more than our own personal experiences. Each person's mind is like an island. While it seems to stand alone, there are subterranean connections via the ocean floor with other islands. The ocean floor can be conceived in a variety of ways. Perhaps it is a "racial unconscious," as Jung would say (1922); perhaps it is the *atman* (the universal soul) of Hindu psychology. Perhaps it is "objective mind" — the *objektiver Geist* of Hegel, the island formations representing so many individuations of one universal psychic system.

William James, though certainly not a Hegelian, employed a similar postulate to account for certain religious experiences, especially the apparent inrush of religious conviction and comfort during moments of prayer or meditation.

> The further limits of our being plunge, it seems to me, into an altogether other dimension of existence from the sensible and merely "understandable" world. Name it the mystical region, or the supernatural region, whichever you choose. (1902, p. 506)

To James this metaphysical assumption did not seem far-fetched. He wrote at a time when the subconscious self had become an accredited phenomenon. Who could tell the limits to its subliminal reach? He felt that the concept was "exactly the mediating term required" to relate the scientific study of the individual to the superindividual verities of religion. In terms of consciousness we are single individuals, but in terms of subconsciousness each of us has a wider self than he knows.

The metaphysical view here represented is not popular in present-day social psychology. Yet there are those who argue that without some similar doctrine it is hard to explain certain common ideas that occur in the race. Jung's contention, for example, is that certain symbols recur, especially in dreams, regardless of specific contact between cultures (1922). For instance, an individual entirely ignorant of Egyptian lore may in his dreams or fantasy employ the scarab as a symbol of immortality. In so doing he dips into a racial unconscious.

Objective mind. We need to look more closely at the Hegelian conception of the *objektiver Geist,* for it has seeped into much sociological and psychological thinking. Brett, in *A History of Psychology* (1921, Vol. 3), argues that we owe to Hegelian philosophy the beginning of social psychology. For Hegel's system is the true antecedent of all theories which pass beyond individuals and claim that in some sense the social mind has unity.

According to Hegel's idealistic philosophy there is only one Mind (1807). It is absolute, all-embracing, divine. It works itself out in the course of history. Individual men are but its agents. Its principal focus is in the state, which is therefore the chief agent of divine life on earth. Each state has, in fact it *is,* a group mind. It has its own laws of growth and development (the dialectic) and while it makes much use of individuals, it is by no means reducible to their transitory mental life. Marx, as well as Hitler, was among the sinister spiritual children of Hegel. Like him, they equated personal freedom with obedience to the group, morality with discipline, personal growth with the prosperity of the party, class, or state. *Du bist nichts: dein Volk ist alles* was the Nazi rallying cry.

We can trace Hegel's psychological influence in several directions. As we have said, it underlay Karl Marx's exaltation of social class as a superindividual entity. All contemporary Soviet psychology reflects this view (Bauer, 1952). In Britain, Bosanquet (1899) and Green (1900) were among the political philosophers who, following Hegel, viewed the state as an organic mind transcending the component minds of individuals, and demanding "sober daily loyalty." It is hardly necessary to point out that psychological apologists for racism and nationalism (e.g., Jaensch, 1938) tend no less than Hegel to apotheosize the group mind, as represented by the state, race, folk, or *Kultur*. Schaeffle (1878), while approaching the matter by way of the organic analogy, completed this analogy by accepting the nation as an organism in the fullest sense of the word, as capable of having its own goals and purposes, its own pleasures and pains, its own consciousness, and its own right to demand obedience from the individual.

Even without subscribing to the moral aspects of Hegelian philosophy, many social psychologists in Germany have taken up his concept of *Volksgeist*. To them it seems self-evident that a people has a mental unity. Hence the goal of social psychology is the discovery of "what must be added to the sum of individuals to account for the creation of a unity out of multiplicity, and what aspects of individual action conform to the action of others so that a harmony of activity results" (Lukasczyk, 1951, p. 51). In the following section we shall consider *Völkerpsychologie*, which especially busied itself with this problem.

In France, Espinas (1877) set forth a theory of an objective group mind, apparently uninfluenced by Hegel. Individual minds, he held, fuse into a "collective consciousness." They do so because presentations (ideas) pass rapidly from mind to mind by means of language and gestures; impulses (emotions) likewise are contagious. Thus it is from the normal processes of communication that a group mind is generated. While individual minds are multiple, yet there is engendered a social consciousness capable of knowing itself, and in which each individual participates.

Folk psychology. In addition to Hegel, several writers, such as Fichte and von Humbolt, laid the philosophical ground for the concept of *Volk* which gradually emerged in the course of Germany's struggle for unification. By 1860 the concept was ripe for psychological exploitation. At that time three programs for research and theory were laid down, all of them aiming to combine ethnological study with psychological interpretations. Bastian (1860) introduced his concept of *Elementargedanken* — thought-forms so basic to the life of a folk that they would supply the unit of analysis needed for a study of any form of human society. Lazarus and Steinthal (1860) inaugurated the *Zeitschrift für Völkerpsychologie und Sprachwissenschaft*, which had an influential life of thirty years, ceasing publication in 1890. Wundt (1862) wrote his remarkable *Beiträge*, in which he laid down a program of writing and research for himself that should (and did) occupy him for sixty years, reaching its culmination in the publication of ten volumes of his *Völkerpsychologie* (1910-1920). While these authors worked from the armchair (or at "the green table," as the Germans say), the movement they represented had a lasting influence on social science. In Germany, Thurnwald (1925-1935), and in America Judd (1926) and Goldenweiser (1933), are at least in part representatives of the tradition.

Some years later England picked up the thread with empirically skillful fingers. The conviction that anthropology needs psychology found its first expression in authentic collaborative field work. In 1898 the Cambridge Anthropological Expedition to Torres Straits enlisted in its field staff three trained experimental psychologists, W. H. R. Rivers, W. McDougall, and C. S. Myers. Here for the first time a study with laboratory equipment was made of preliterate peoples under their ordinary conditions of life. The expedition was a landmark in the collaboration of social scientists. The spirit of this historic research is still alive and expanding. Today ethnologists often borrow both theory and methods from psychology, and psychologists in turn are growing more and more aware of the need for testing their hypotheses cross-culturally. In this connection one thinks of the work of Freud (1913), Bartlett (1923), Róheim (1925), Rivers (1926), Kluckhohn (1944), Malinowski (1944), to mention only a few. It is not difficult to account for the interdependence of social psychology and ethnology (cultural anthropology). Taken by itself, ethnology lacks a theory of the motives and capacities of men; whereas social psychology by itself cannot state the cultural models to which men's motives and capacities tend to conform.

Lazarus and Steinthal were a collaborative pair, an anthropologist and a philologist, primarily interested in a comparative characterology of peoples. Each people, they held, had its own *Volksgeist*. This *Geist* they defined in the opening volume of their Journal (1860) as "a similar consciousness of many individuals, plus an awareness of this similarity, arising through similar descent and spatial proximity." In this definition we may, if we wish, see merely a mild nominalism. The reality of the group mind is not transcendental, it is only a matter of similarities of individual minds plus a recognition of these similarities. But, unfortunately, this easily intelligible concept is not consistently maintained. In most of their discourse Lazarus and Steinthal speak of the *Volksgeist* in true Hegelian fashion as a metaphysical entity standing over against the individual mind. This type of double-talk was apparently encouraged by the influence of Herbart, whom the authors followed. Herbart had written an essay proposing that the static and dynamic conditions of the individual mind had parallels in the statics and dynamics of the state (1821). Just as ideas might rise above, or sink below, the threshold of an individual's consciousness, so too they may grow active or fall into disuse in society. Just as ideas may conflict in an individual's mind, so too may there be a war of ideas in society. Strictly speaking, Herbart's own position is merely analogical, but it was interpreted as establishing an independent self-active *Geist* of the people.

The embodiment of the folk mind lies in speech, mythology, religion, folklore, art, literature, morality, custom, and law. Hence the appropriate methods for its study are those familiar to ethnologists and to philologists. With this starting point, it is not surprising that little progress was made in developing a social psychology, as distinct from traditional ethnology. In spite of programmatic fanfare, psychological analysis did not advance. When, on rare occasions, psychology was invoked it was invariably the idea-psychology of Herbart.

Wilhelm Wundt (1832-1921), in spite of his controversies with Lazarus and Steinthal, was much like them in his conception of *Völkerpsychologie*. The chief systematic difference between them was that Wundt was not a Herbartian but a Wundtian. Having evolved his own fundamental concepts in general psychology, he drew upon them freely in his social psychology. Wundt decided early in his career (1862) that there were two branches of psychology — physiological and social (folk). He planned to devote the first part of his life to the former subject, and the second part to the latter. He kept to his plan, writing prodigiously — an average of two and two-tenths pages per day for a span of sixty-eight years, 53,735 pages in all (Boring, 1950, p. 345).

The most significant fact about Wundt is his insistence that the study of all higher mental processes falls in the province of *Völkerpsychologie*. He did not believe that individual psychology, especially as pursued in the psychological laboratory, could account for men's thought. Thinking is heavily conditioned by language, by custom, and by myth, which to him were the three primary problem areas of *Völkerpsychologie*. The argument runs as follows: when men receive sensations from the outer world these commence to combine according to the laws of association, but association depends upon the absorption of an impression into the individual's apperceptive mass. Impressions enter into a creative relation with memories and contexts stored in this apperceptive mass. But this apperceptive mass is itself largely a product of culture. The apperceptive mass is furnished with linguistic habits, moral ideas, and ideological convictions that compose the *Volksseele*. Wundt would have felt at home in modern-day discussions of "social perception."

He preferred the term *Volksseele* to *Volksgeist* which, he insisted, was too objective, implying an entity entirely apart from, and over and above, individuals. Semantically at least, Wundt leaned somewhat away from the substantive group mind implied by Hegel and by Lazarus and Steinthal. Yet, like all group-mind theorists, Wundt seems to temporize. He says that while it is true that the folk soul presupposes component individuals, yet it is more than the sum of their mental activities. Interaction brings about new qualities. A language, for example, is a psychic product and a determining force that cannot possibly be accounted for excepting in terms of interaction. Josiah Royce was following Wundt when he wrote:

> The creator of the English speech is the English people. Hence the English people is itself some sort of mental unit with a mind of its own. (1913, p. 27)

The case for the folk soul rests to a large extent upon Wundt's concept of *psychic actuality*. Although in a broad way Wundt subscribed to psycho-physical parallelism, yet he believed that mental events can be regarded as lying in a causal chain. Men's experience is not identical with the activity of nerves. One state of consciousness may lead directly to another, and one man's conscious life may interact with another's. Thus by employing a somewhat loose hypothesis of the body-mind relationship Wundt is able, without serious difficulty, to claim a kind of autonomy for *group* mental life.

We have said that, to Wundt, the primary ingredients of the *Volksseele* were language, myth, and custom. By the time he had exhausted these topics in his ten-volume *Völkerpsychologie* he had, however, embraced art, religion, law, and social organization. But he omits altogether the material and technological aspects of culture. He found no use for the skulls, tools, or potsherds that fascinate anthropologists. Why? Because as a psychologist Wundt identified the mind of a folk only with its immaterial culture, i.e., with the habits and ideas that lead to activity, not with the *products* of activity.

Summary. The movement we have been discussing is a curious blend of Hegelian idealism, Herbartian idea-psychology, Wundtian concepts, and library ethnology. While its central intent was to discover the nature and laws of the group mind, the actual accomplishment in theory was meager and confused. The movement did, however, succeed in preparing the ground for fruitful collaboration between psychology and social anthropology. Today this tide of collaboration is still rising, and has not yet reached its crest. An important clarification of concepts has occurred, as Kroeber and Kluckhohn (1952) point out. What was formerly ascribed to the mind of the *Volk* is now largely subsumed under the current concept of *culture.* "Personality in culture" promises to be a far more productive concept than "the mind of the group" has been.

Collective representations. Emile Durkheim, 1858-1917, was a French contemporary of Wundt. In many ways the views of the two authors were similar. Both felt keenly "the reality of the social." Both insisted that individual minds were in some sense compounded into a collective mind. Each had a favorite example of the force that this collective mind may exert upon the individual: Wundt chose language, Durkheim religion. Both tended toward intellectualism, regarding the social mind as furnished with ideas, representations, apperceptions (rather than with emotions, passions, volitions). Of the two, however, Wundt was the more voluntaristic.

In one respect, however, Durkheim was more extreme than Wundt. He rejected the parallelistic view that men's bodies and men's minds are somehow closely bound together. Wundt, with his concept of "psychic actuality," had taken liberties with this doctrine perhaps more than he knew. But Durkheim repudiated it altogether. To him events in the nervous system were one thing, events in the individual mind something entirely different, and events in the social mind a third distinctive type of reality.

His argument, as set forth in the *Révue de métaphysique* (1898) is trenchant. He first sets out to demolish the epiphenomenal view that images, memories, thoughts are merely the excrescence of brain activity. William James had identified memory with brain traces, and denied that memory was a fact of the mental order at all. Huxley had likened thought to the shadow cast by the machinery of the brain. Durkheim ridicules the claim. If mental states are only an echo of the physical, why do we have conscious mental states at all? And this was a question that perplexed William James as well, and drove him, ultimately and inconsistently, to argue as Durkheim did for a certain independence of the psyche.

Durkheim was willing to admit that only sensations, the primary elements of individual states of mind, are referable directly to cerebral states. But once existing, we find that sensations compound themselves by the laws of association into images and ideas. These are no longer referable to brain states. Why not? For one thing the basic law of association is similarity (similar ideas form larger clusters or concepts), and there is no mechanism in the brain for producing similarity. While the point is too technical to discuss here, Durkheim's attack on the identical-element theory of similarity is still cogent. It is only whole ideas, he argues, that are similar to one another; they are not decomposable into elements or reducible to brain states.

Having established to his satisfaction that individual representations (ideas) are not reducible to nervous operation, he then continues his case for emergence.

> Just as you cannot specify the contribution of each brain cell to an image, so you cannot specify the contribution of an individual to the collective representation. (1898, p. 296)

Thus having leaped the physiological-psychological gulf, he then leaps the psychological-social gulf. While individuals contribute to the *conscience collective,* this final emergent is exterior to, and independent of, any specified individual:

> When we have said that social facts are in a sense independent of individuals and exterior to individual consciousness, we have only affirmed in the social realm what we have established in the psychic realm. (p. 293)

While higher levels emerge from lower, there is a new autonomy achieved at each stage.

> brain cells → sensations → images → individual representations → collective representations.

Not only is Durkheim a dualist in his separation of mind and body, but from the point of view of monists, he compounds the felony by arguing for the independence of the social and the individual. There are two consciousnesses in each of us, he insists: one pertains to our own private experience, the other to all portions of experience attributable to the fact of human association. This association of human beings produces thoughts and manners and practices that are external to, and independent of, any single human being. Religion (1912) is a collective representation to which the individual in his course of growth is exposed. The division of labor (1902) is similarly a fact of the social mind, not of the individual. Social disruption (*anomie*), even the causes of suicide (1897), are exterior to any one member of the group.

Besides having the property of *exteriority,* collective representations have likewise the property of *constraint.* There is a compelling force in society that requires the individual to behave according to its rules. Often, of course, we cooperate and conform gladly. The concept of constraint covers both cooperation and coercion.

Durkheim, it is true, is a sociologist, yet the line of reasoning by which he achieves his "sociological sociology" is of special interest to the psychologist. His argument is more painstaking, more sophisticated, than that of most writers who have postulated a group mind. There are many secondary sources that give fuller treatment of Durkheim's contributions, among them Gehlke (1915), Goldenweiser (1933), Parsons (1937). Durkheim's influence upon anthropological theory may be seen strikingly in Lévy-Bruhl's work, *La mentalité primitive* (1922).

While the concept *collective representations* is peculiarly Durkheim's, other authors — psychologists among them — have shown its influence. Take, for example, Fouillée's key concept of the *idées-forces* (1908). While this writer believed that Durkheim had pushed the criteria of exteriority and constraint too far, yet he accepted the basic fact that much of each person's mental life is of social origin. He proposed that the so-called constraint of social ideas lies in the psychologically anchored mechanism of ideo-motor action. Fouillée saw no need to break with psychology as drastically as Durkheim had done.

The latter's influence may be traced also in Piaget (1932), who regards the *moral realism* of the child as the result of the child's exposure to, and unquestioning acceptance of, prevailing conceptions of right and wrong, good and bad. Similar is Bartlett (1932), who regards memory as heavily conditioned by the thought forms current in a social group. While Piaget and Bartlett would not subscribe to Durkheim's view as a whole, the problems they treat and their manner of approach show a Durkheimian respect for the reality of the social.

Cultural determinism. Strictly speaking, cultural determinism is not a theory of the group mind. It is, however, a doctrine that makes much of the exteriority and constraint of cultural factors. Extreme advocates seem at times to be saying that psychology can explain nothing of importance concerning social behavior. They quote with favor Durkheim's adage, "Every time a social phenomenon is directly explained by a psychological phenomenon, we may be sure that the explanation is false." The following passage, written by White (1947), presents the argument:

> To the culturologist the reasoning that says that one people drinks milk because they like it, another do not because they loathe it, is imbecilic. . . . It explains nothing at all. Why does one people like, another loathe, milk? This is what we want to know. And the psychologist cannot give us the answer. Nor can he tell us why a people does or does not avoid mothers-in-law, practice monogamy, inhumation, the couvade, or circumcision; use chopsticks, forks, the pentatonic scale, hats, or microscopes; form plurals by affixation — or any of the other thousands of customs known to ethnography.

White favors a *cultural* explanation even for the existence of nonconformists:

> Even a revolutionist is not an individualist: he is a human organism that is held and wielded by certain cultural elements that are moving in the direction of profound change.

This manner of thinking about social behavior is at a high level of abstraction. Only gross changes, broad trends, and long-range effects are considered. Culturologists happen to be interested in wide horizons; to them the individual is unimportant. They do not worry about the fact that all cultures have plenty of men who deviate from the picture they create of a "modal personality." Even in a milk-drinking culture, plenty of individuals loathe milk. Nor do culturologists worry about the extreme selectivity individuals show in responding to this or that feature of their own culture. Nor are they concerned with the process of communication and learning by which children within a culture acquire its ways. Cultural determinism says in effect that children in English-speaking countries will learn to speak English, but it does not explain why they learn to speak at all, or how they learn, or why among them there are almost infinite variations in habits of speech.

The quarrel, of course, is unprofitable. Let us admit that cultural models prescribe what types of conformity will (probably) be learned. Also that over long periods of time language and other cultural forms seem to go through cycles of change independently of what any particular individual may do or think. Yet within this cultural matrix certain psychological factors are presupposed: motivation, learning, the processes of perception, concept formation, and the organization of attitudes and sentiments. Only when the cultural and the psychological factors are treated together is a full explanation of social behavior likely to be achieved.

Common and reciprocal segments of behavior. We have pointed out that the term "group mind" disappeared from the scene after the 1920's, even though the problem itself has persisted. Boredom had much to do with dropping the label. Everyone seemed wearied by the climax of the controversy in the decade of the twenties.

Certain nominalists attacked the "group fallacy" with vigor. Thus F. H. Allport (1924) wrote:

> Nationality, Free-masonry, Catholicism, and the like are not group minds . . . they are sets of ideals, thoughts, and habits repeated in each individual mind and existing only in those minds. . . . All theories which partake of the group fallacy have the unfortunate consequence of diverting attention from the true locus of cause and effect, namely, the behavior mechanism of the individual. (p. 9)

This same author expanded and fortified his argument in *Institutional Behavior* (1933). Opposing views, based largely on the argument of cultural determinism, were stated by Wallis (1925) and Judd (1926). A judicious weighing of evidence, pro and con, came from the pens of the philosophers Perry (1922) and Hocking (1926).

Best known among the disputants is McDougall (1920). The first edition of his *Group Mind* seemed to affirm, with qualifications, his belief in its existence; the second edition (1928) bewailed the reader's misunderstanding of his position. Earlier in this section we stated McDougall's criteria for the existence of a group mind. He found these fulfilled most clearly in the case of nations, for nations have a permanent and highly organized character. McDougall does not regard his criteria as leading to the postulation of a social consciousness, for "there is no consciousness excepting in individuals"; yet, he insists, "we may still speak of collective minds, for we have defined mind as an organized system of interacting mental or psychical forces" (p. 66).

In spite of their disagreements, there is a common thread in the thinking of McDougall and of F. H. Allport. What is called the group mind, they say, is essentially the abstraction of certain attitudes and beliefs from the personal mental life of individuals. We may, if we wish, regard all the common (comparable) attitudes of Englishmen as comprising the group mind of the English nation. McDougall, in the Preface to the 1928 edition, explicitly states that he views the group mind

> as consisting in the similarities of structure of the individual minds which render them capable of responding in similar fashion to the common features of the environment, social and physical. (p. ix)

Social institutions, thus regarded, become abstractions from the behavior and consciousness of individuals. People are not parts of institutions, but institutions are parts of people. Figure 3 represents the situation diagrammatically.

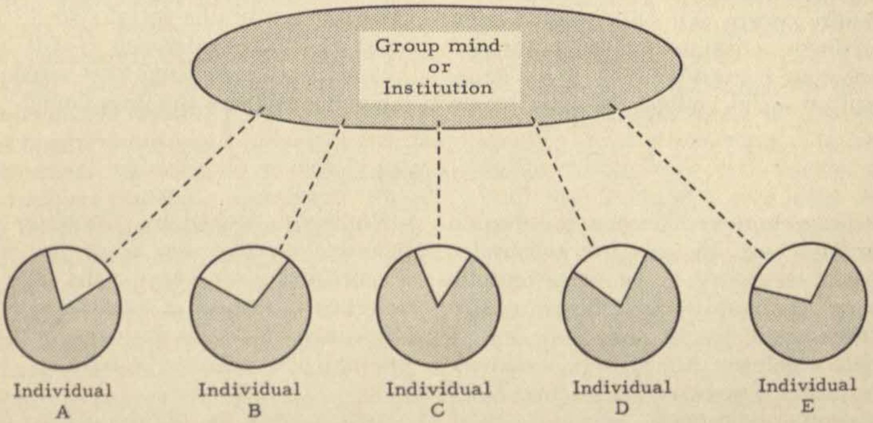


FIG. 3. Common and reciprocal segment theory.

It should be noted in Fig. 3 that *common* segments (i.e., similar habits and attitudes) are clearly represented, but that the *reciprocal relation* of these segments — also essential to a group mind or institution — is not well represented. As it stands, the diagram overstresses *conformity,* which is indeed an important aspect of institutionalism, but understresses the *reciprocity of roles,* which is an equally important aspect. Both McDougall and F. H. Allport (1940) admit that both aspects are vital. The latter author considers conformity and reciprocity to be the essential features of operationally denotable "event structures" that have in the past been called group minds or institutions.

Such a view preserves the strictly individual locus of mind, yet enables us for certain purposes to examine segments of the mental life of individuals in relation to comparable, or reciprocal, segments of the mental life of others. This method permits culturologists to affirm the importance of custom, viewed as common conformity. It allows the sociologist to affirm the independence (in a sense) of institutions, the activity of social systems, and the viability of role relations.

Final word. After McDougall had wrestled with the problem for a time, he bewailed his "tactical error" in using the phrase "group mind." Probably it is regrettable that the concept was ever used by anyone. We see now that it has unnecessarily imposed metaphysical blocks in the path of constructive conceptualization. It is important, however, for the student of social science to understand this historical background. Corporate groups have been multiplying at a spectacular rate throughout modern society. An individual is a member of many publics, of many institutions, of many social systems. Each of these memberships poses afresh the ancient question. The problem of the group mind will not be downed.

But at the present time theories are being recentered. Writers realize that the integrity of the personality system can be preserved even while we affirm the existence and action of social structures that are transindividual. It was Sapir who advised all social and psychological scientists to form the habit of looking at their data both from the concrete individual point of view and from the abstract social point of view. It enriches research and theory to do so. As we have said, the approach now growing in favor combines the idea of "common segment" with that of "role." Parsons and Shils (1951) write:

> Personality as a system has a fundamental and stable point of reference, the acting organism. It is organized around the one organism and its life processes. But ego and alter in interaction with each other also constitute a system. This is a system of a new order, which, however intimately dependent on them, does not simply consist of the personalities of the two members. (p. 23)

Having thus restated the problem in terms of "systems," the authors proceed to say that "the most significant unit of social structures is not the person but the role." Both the actor and those with whom he interacts possess certain expectations due to their places in the social structure. The final essential step in the argument, obviously akin to the common segment theories, is stated as follows:

> The abstraction of an actor's role from the total system of his personality makes it possible to analyze the articulation of personality with the organization of social systems. The structure of a social system and the functional imperatives for its operation and survival or orderly change as a system are moreover different from those of personality. (*loc. cit.*)

Thus what was formerly called the "group mind" becomes a matter of abstracted role segments, both common and reciprocal. These segments or aspects of personality provide the raw material for the social system. While this mode of conceptualization is still incomplete, it provides a promising new attack upon the venerable riddle of the group mind.

UNITS OF ANALYSIS

A firm science employs definitely agreed upon units of analysis: e.g., the quantum of physics, the element of chemistry, or the nerve impulse in physiology. By contrast, psychology has successively tried out (and in part discarded) a wide variety of elusive units: e.g., sensation, image, idea, reflex, S-R bond. Social psychology has had a similar experience. It would like to have some relatively stable unit (or units) of behavior out of which the theorist can compound a systematic view of social interaction. But, as in the case of general psychology, the search has met with frustrations. Some social psychologists, especially Gestaltists and field theorists, say that the search for elements is doomed to failure because the individual does not react to stimuli in fixed ways; he alters his adjustment to accord with the total pressure of internal and external conditions (e.g., Asch, 1952).

Still, historically, the postulation of units has played an important part in the development of social psychology, and it is by no means clear that they can be, or should be, dispensed with. Most important among the assumed units are *instinct, habit, attitude, sentiment.* From this list we shall leave out *idea,* which, as we have already made clear, declined in popularity at the end of the nineteenth century.

Instinct. While faculty psychologists since the time of Descartes affirmed and classified the "passions" of men, their efforts led to no systematic discussion of the nature of these faculties or of their social consequences. Darwin's doctrine of natural selection changed the situation. It gave grounds for believing that all behavior — animal and human, individual and social — lay in the evolutionary array of instincts that make for survival of the species. Following Darwin's lead, naturalists immediately began to prepare lists of animal instincts. It was not long before writers of psychological textbooks (e.g., James, 1890) set forth with more abandon than reflection long schedules of the presumably basic motivational units underlying human behavior (*cf.* Bernard, 1926). But it remained for McDougall to clarify the implications of Darwinism and build a completely coherent system of social psychology upon the instinct hypothesis. His definition of instinct, if fully understood word by word, adumbrates his entire system of social psychology.

> We may, then, define an instinct as an inherited or innate psycho-physical disposition which determines its possessor to perceive, and to pay attention to, objects of a certain class, to experience an emotional excitement of a particular quality upon perceiving such an object, and to act in regard to it in a particular manner, or, at least, to experience an impulse to such action. (1908, p. 30)

Under the broad tent of this definition McDougall is able to order many, if not most, social phenomena. The process of suggestion is a consequence of the instinct of self-abasement. Primitive passive sympathy, as we have seen, is a matter of induced instinctive seizure when one person sees another behaving in an instinctive manner. Laughter is an antidote to such sympathetic induction. Emotion is lockstitched to instinct. Instincts may form compounds: thus a person's religious orientation may be a blend of curiosity, self-abasement, flight (fear), and the tender emotion which accompanies the parental instinct. McDougall believed that the whole story of social behavior could be based on his hormic (urge-like) view of instincts. He readily admitted that the force of learning and integration creates complex units (sentiments) that are often of more immediate concern than instincts to social psychologists. But sentiments, he insisted, are not the ultimate units of behavior. Only instincts are ultimate.

The decade following the appearance of McDougall's *Social Psychology* was almost wholly dominated by instinct theory. We have already called attention to G. Wallas' *Human Nature and Politics* (1908), Trotter's *Instincts of the Herd in Peace and War* (1916). We may add Thorndike's *The Original Nature of Man* (1913), Woodworth's *Dynamic Psychology* (1918), and Dewey's presidential address to the American Psychological Association, in which he declared that the science of social psychology must be founded on a doctrine of instincts (1917).

The same decade was marked by the world-wide spread of Freud's theories. While Freud's instinctivism differs in many ways from McDougall's, it too helped fix the attention of social psychologists upon the primacy of impulses in social conduct. During the decade 1910-1920, however, Freud's influence on *social* psychology was not as marked as McDougall's, probably because the instincts Freud postulated were far less definite and manipulable than McDougall's. By 1932 McDougall had given up the term "instinct" (in response to violent attack by behaviorists), but maintained all essential features of his doctrine under the label "propensities." Between 1908 and 1932 he had considerably expanded his list of native propensities. Its final form includes eighteen items:

> Food-seeking; disgust; sex; fear; curiosity; protective or parental propensity; gregariousness; self-assertion; submission; anger; appeal; constructive propensity; acquisitive propensity; laughter; comfort; rest or sleep; migratory propensity; a cluster of specific bodily needs — coughing, sneezing, breathing, elimination. (McDougall, 1932, Chap. 7)

In 1919 Knight Dunlap hurled the first anti-instinct bombshell. From that time onward the whole theory came under attack. Objections to it are many. For one thing, there is no factual proof for the existence of instincts in McDougall's sense of the term. For another, the doctrine brings with it a curious sterility in theorizing. It implies that what is given by heredity is final. Having discovered "first causes" and "prime movers" there is no incentive to look further and discover additional basic principles of human behavior. Furthermore, instinct doctrines inevitably regard certain motives as primary, and all others as secondary, derived, and less important. It is questionable whether such a distinction can be sustained. McDougall himself felt compelled to supplement instincts by sentiments, but do such complex forms of moral and socialized behavior merely borrow their energy from a handful of original and unchanging instincts? Finally, American social scientists generally have a distaste for nativism, the view that human nature is fixed and at bottom unalterable. The American intellectual climate favors environmentalism.

But it must not be thought that instincts were discarded altogether. How could they be? Elementary motives and reflex adjustments are manifestly present in infancy prior to training and independent of experience. Even parsimonious behaviorists had to admit innate pushes of some sort, variously called "tissue changes," "prepotent reflexes," or "primary drives." But these initial motives were not viewed, as in McDougall's case, as enduring and changeless units of human nature; they became incorporated through learning into a chain of *habits* which constituted the "true" unit.

Besides behaviorism, which represents a distinct countertrend to instinctivism, there is a somewhat colorless middle position that many authors today prefer to hold. They, like McDougall, feel the necessity for dynamic units, and variously affirm the existence of *desires, wishes, needs, vectors,* or simply, *motives.* Such units are employed as building blocks in various contemporary systems, but they differ from earlier instinctive units in that they try to sidestep the thorny question of innateness. Modern writers seem not to care whether a motivational unit is inborn or learned. Likewise, they tend to hold their postulated units in a tentative, heuristic way — as mere "constructs" — and not to regard them as eternally fixed in the person or in the species.

Habit. Although James affirmed the existence and importance of instinct, he gave chief weight to the habit unit. Instincts as such never endure; strictly speaking, they manifest themselves only once in a lifetime, thereafter learning immediately sets in. This doctrine of the *transitoriness of instinct* prepared the way for James's famous apotheosis of habit. It is more than second nature, he insisted, it is "ten times nature."

> Habit is thus the enormous flywheel of society, its most precious conservative agent. It alone is what keeps us all within the bounds of ordinance, and saves the children of fortune from the envious uprisings of the poor. . . . It keeps different social strata from mixing. Already at the age of twenty-five you see professional mannerisms settling down on the young commercial traveller, on the young doctor, on the young minister, on the young counsellor-at-law. . . . It is well for the world that in most of us, by the age of thirty, the character has set like plaster, and will never soften again. (1890, I, p. 121)

The social and ethical consequences of this view of habit are, according to James, "numerous and momentous." Since in his view we acquire habits primarily through repetition of actions, it behooves us to pay alert attention to the formation of right habits. In building good habits never let an exception occur. Whether they are good or bad, all habits are a ball-and-chain.

It was partly due to the eloquence of James that the habit unit came to be taken for granted during the following three or four decades. To be sure, it became recast in terms of "conditioned response" in the hands of behaviorists, but this unit is only another version of the habit doctrine: Holt (1915), Watson (1919), F. H. Allport (1924). Somewhat later, in conjunction with what is sometimes known as "Yale learning theory," the concepts "response" and "habit family hierarchy" became preferred (e.g., Miller and Dollard, 1941). But while the habit theory has grown more sophisticated than that depicted by James, the unit is of the same genre.

A somewhat different conception of habit was employed systematically by Dewey "as the key to social psychology" (1922), only a few years after he forsook his advocacy of the instinct unit. In two important respects Dewey's view differs from most others. In the first place, the unit is more flexible; it is not merely a frozen connection in motor pathways, it includes general attitudes, outlooks, purposes, interests. A man may commit murder only once in his lifetime. The act is, however, due to habit — to a habit of hatred. Furthermore, habit has a motivational character:

> We may think of habits as means, waiting, like tools in a box, to be used by conscious resolve. But they are something more than that. They are active means, means that project themselves, energetic and dominating ways of acting. (p. 25)

In the second place, Dewey tried to include situationism within his conception of habit. The neural mechanism alone does not anchor the habit. The environment helps to sustain it. If the environment changes the habit will change. Every manifestation of a habit is altered by other people who are concerned in the act. "Conduct is always shared; this is the difference between it and the physiological process" (p. 17).

Dewey was attempting to claim too many attributes for the habit unit (fixity and change, neural impulse and situation, motivation and instrumentality). As a result, the unit seemed vague. Later writers did not follow his lead or attempt to refine the concept for systematic use. The reason, no doubt, is that the concept *atti-*

tude was already established. For all intents and purposes it filled the need that Dewey felt.

Attitudes. The following brief historical survey of the attitude concept is abbreviated from the author's earlier and more extended chapter on "Attitudes" published in *The Handbook of Social Psychology,* edited by C. C. Murchison (1935).

This concept is probably the most distinctive and indispensable concept in contemporary American social psychology. No other term appears more frequently in experimental and theoretical literature. Its popularity is not difficult to explain. It has come into favor, first of all, because it is not the property of any one psychological school of thought, and therefore serves admirably the purposes of eclectic writers. Furthermore, it is a concept which escapes the controversy concerning the relative influence of heredity and environment. Since an attitude may combine both instinct and habit in any proportion, it avoids the extreme commitments of both the instinct theory and environmentalism. The term likewise is elastic enough to apply either to the dispositions of single, isolated individuals or to broad patterns of culture (common attitudes). Psychologists and sociologists therefore find in it a meeting point for discussion and research. This useful, one might almost say peaceful, concept has been so widely adopted that it has virtually established itself as the keystone in the edifice of American social psychology. In fact, several writers, starting with Thomas and Znaniecki (1918), have *defined* social psychology as "the scientific study of attitudes."

Like most abstract terms in the English language, *attitude* has more than one meaning. Derived from the Latin *aptus,* it has on the one hand the significance of "fitness" or "adaptedness," connoting, as does its by-form *aptitude,* a subjective or mental state of preparation for action. Through its use in the field of art, however, the term came to have a quite independent meaning; it referred to the outward or visible posture (the bodily position) of a figure in statuary or painting. The first meaning is clearly preserved in the phrase "mental attitudes," and the second meaning in "motor attitudes." Since mentalistic psychology historically precedes response psychology, it is only natural to find that mental attitudes are given recognition earlier than motor attitudes. One of the earliest psychologists to employ the term was Herbert Spencer. In his *First Principles* (1862), he wrote:

> Arriving at correct judgements on disputed questions, much depends on the attitude of mind we preserve while listening to, or taking part, in the controversy; and for the preservation of a right attitude it is needful that we should learn how true, and yet how untrue, are average human beliefs. (Vol. I, 1, i)

Somewhat later, when psychologists were forsaking their exclusively mentalistic point of view, the concept of *motor attitudes* became popular. In 1888, for example, N. Lange developed a motor theory wherein the process of a perception was considered to be in large part a consequence of muscular preparation or "set." At about the same time Münsterberg (1889) developed his action theory of attention, and Féré (1890) maintained that a balanced condition of tension in the muscles was a determining condition of selective consciousness.

In recent years it is uncommon to find explicit labeling of an attitude as either "mental" or "motor." Such a practice smacks of body-mind dualism, and is therefore distasteful to contemporary psychologists. In nearly all cases today the term appears without a qualifying adjective, and implicitly retains both its original meanings: a mental aptness and a motor set. Attitude connotes *a neuropsychic state of readiness for mental and physical activity.*

Perhaps the first explicit recognition of attitudes within the domain of laboratory psychology was in connection with a study of reaction time. In 1888 L. Lange discovered that a subject who was consciously prepared to press a telegraph key immediately upon receiving a signal reacted more quickly than did one whose attention was directed mainly to the incoming stimulus, and whose consciousness was not therefore directed primarily upon the expected reaction. After Lange's work, the task attitude, or *Aufgabe,* as it came to be called, was discovered to play a decisive part in nearly all psychological experiments. Not only in the reaction experiment, but in investigations of perception, recall, judgment, thought, and volition, the central importance of the subjects' *preparedness* became universally recognized. In Germany, where most of the early experimental work was done, there arose a swarm of technical expressions to designate the varieties of mental and motor "sets" that influence the subjects' trains of thought or behavior during the experiment. In addition to the *Aufgabe,* there was the *Absicht* (conscious purpose), the *Zielvorstellung* (or idea of the goal), the *Bezugsvorstellung* (idea of the relation between the self and the object to which the self is responding), the *Richtungsvorstellung* (or idea of direction), the *determinierende Tendenz* (any disposition that brings in its train the spontaneous appearance of a determined idea), the *Einstellung,* a more general term (roughly equivalent to "set"), the *Haltung* (with a more behavioral connotation), and the *Bewusstseinslage* (the "posture or lay of consciousness"). It was perhaps the lack of a general term equivalent to "attitude" that led the German experimentalists to discover so many types and forms. The lack may also explain why no systematic social psychology was written in Germany based upon a unified concept of "attitude."

Then came the lively controversy over the place of attitudes in consciousness. The *Würzburg* school was agreed that attitudes were neither sensation, nor imagery, nor affection, nor any combination of these states. Time and again attitudes were studied by the method of introspection, always with meager results. Often an attitude seemed to have no representation in consciousness other than a vague sense of need, or some indefinite and unanalyzable feeling of doubt, assent, conviction, effort, or familiarity (Titchener, 1909).

As a result of the *Würzburg* work all psychologists came to accept attitudes, but not all believed them to be impalpable and irreducible mental elements. In general, the followers of Wundt believed that attitudes could be accounted for adequately as *feelings,* particularly as some blend of striving and excitement. Clarke (1911), a pupil of Titchener, found that attitudes in large part *are* represented in consciousness through imagery, sensation, and affection, and that where no such states are reported there is presumably merely a decay or abbreviation of these same constituents.

However they might disagree upon the nature of attitudes as they appear in consciousness, all investigators, even the most orthodox, came to admit attitudes as an indispensable part of their psychological armamentarium. Titchener is a case in point. His *Outline of Psychology* in 1899 contained no reference to attitude; ten years later, in his *Textbook of Psychology,* several pages are given to the subject, and its systematic importance is fully recognized.

Many authors have reduced the phenomena of perception, judgment, memory, learning, and thought largely to the operation of attitudes (*cf.,* e.g., Ach, 1905; Bartlett, 1932). Without guiding attitudes the individual is confused and baffled. Some kind of preparation is essential before he can make a satisfactory observation, pass suitable judgment, or make any but the most primitive reflex type of response. Attitudes determine for each individual what he will see and hear, what he will think and what he will do. To borrow a phrase from William James, they "engender meaning upon the world"; they draw lines about, and segregate, an otherwise chaotic environment; they are our methods for finding our way about in an ambiguous universe. It is especially when the stimulus is not of great intensity nor closely bound with some reflex or automatic response that attitudes play a decisive role in the determination of meaning and of behavior.

The meagerness with which attitudes are represented in consciousness resulted in a tendency to regard them as manifestations of brain activity or of the unconscious mind. The persistence of attitudes that are totally unconscious was demonstrated by Müller and Pilzecker (1900), who called the phenomenon "perseveration." The tendency of the subject to slip into some frame of mind peculiar to himself led Koffka (1912) to postulate "latent attitudes." Washburn (1916) characterized attitudes as "static movement systems" within the organs of the body and the brain. Other writers, still more physiologically inclined, subsumed attitudes under neurological rubrics: traces, neurograms, incitograms, brain patterns, and the like. The contribution of the Würzburger and of all other experimental psychologists was in effect the demonstration that the concept of attitude is indispensable.

But it was the influence of Freud that endowed attitudes with vitality, equating them with longing, hatred and love, with passion and prejudice, in short, with the onrushing stream of unconscious life. Without the painstaking labors of the experimentalists, attitudes would not today be an established concept in the field of psychology; but also without the influence of psychoanalytic theory they would certainly have remained relatively lifeless, and would not have been of much assistance to social psychology. For the explanation of prejudice, loyalty, patriotism, crowd behavior, control by propaganda, no anemic conception of attitudes will suffice.

As we have said, the instinct hypothesis did not satisfy social scientists for long, for the very nature of their work forced them to recognize the importance of custom and environment in shaping social behavior. The instinct hypothesis has precisely the contrary emphasis. What social scientists required was a new psychological concept that would escape on the one hand from the hollow impersonality of "custom" and "social force," and on the other from nativism. They gradually adopted the concept of *attitude.*

The credit for instituting the concept as a permanent and central feature in sociological writing must be assigned to Thomas and Znaniecki (1918), who gave it systematic priority in their monumental study of Polish peasants. Before this time the term had made only sporadic appearances in sociological literature, but immediately afterward it was adopted with enthusiasm by scores of writers.

According to Thomas and Znaniecki, the study of attitudes is *par excellence* the field of social psychology. Attitudes are individual mental processes which determine both the actual and potential responses of each person in the social world. Since an attitude is always directed toward some object, it may be defined as a "state of mind of the individual toward a value." Values are usually social in nature, that is to say, they are objects of common regard on the part of socialized men. Love of money, desire for fame, hatred of foreigners, respect for a scientific doctrine are typical attitudes. It follows that money, fame, foreigners, and a scientific theory are all values. A social value is defined as "any datum having an empirical content accessible to the members of some social group and a meaning with regard to which it is or may be an object of activity." There are, to be sure, numerous attitudes corresponding to every social value; there are, for example, many attitudes regarding the church or the state. There are also numerous possible values for any single attitude. The iconoclast may direct his attacks quite at random upon all established social values, or the Philistine may accept them all uncritically. Hence in the social world, as studied by the sociologist, both values and attitudes must have a place.

Following closely in the same vein of thought, Faris (1925) proposed additional refinements. He would distinguish between conscious and unconscious attitudes, between mental and motor attitudes, between individual and group attitudes, and between latent and kinetic attitudes. Park (see Young, 1931), who is likewise in essential agreement with this school of thought, suggests four criteria for an attitude: (1) it must have definite orientation in the world of objects (or values), and in this respect differ from simple and conditioned reflexes; (2) it must not be an altogether automatic and routine type of conduct, but must display some tension even when latent; (3) it varies in intensity, sometimes being regnant, sometimes relatively ineffective; (4) it is rooted in experience, and therefore is not simply a social instinct.

The following are typical definitions of attitude that emerged:

> Attitude = the specific mental disposition toward an incoming (or arising) experience, whereby that experience is modified; or, a condition of readiness for a certain type of activity. (Warren, *Dictionary of Psychology,* 1934)

> An attitude is a mental disposition of the human individual to act for or against a definite object. (Droba, 1933)

> An attitude is a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual's response to all objects and situations with which it is related. (G. W. Allport, 1935)

As we have said, the attitude unit has been the primary building stone in the edifice of social psychology. It has, of course, had many critics. Especially in recent years, learning theorists, field theorists, phenomenologists have attempted to dislodge it. But it is questionable whether their combined efforts can do more than refine the concept for future use. After all, social behavior reflects so much organization, recurrent and habitual expectancy, that the doctrine of attitude (or some close equivalent) is necessary. Without some such concept, social psychologists could not work in the fields of public opinion, national character, or institutional behavior — to mention only a few areas; nor could they characterize the mental organization of social man. The term itself may not be indispensable, but what it stands for is. McDougall has argued that "tendency" is the basic postulate of all psychology (1937). While he did not like the term "attitude," he was affirming the need for some corresponding unit of analysis. His preference was for "sentiment."

Sentiment. According to McDougall:

> The theory of sentiments is the theory of the progressive organization of the propensities (instincts) in systems which become the main sources of all our activities; systems which give consistency, continuity and order to our life of striving and emotion; systems which in turn become organized in larger systems, and which, when harmoniously organized in one comprehensive system, constitute what we properly call *character.* (1932, p. 211)

Thus conceived sentiment differs from attitude in four respects. (1) A sentiment presupposes underlying propensities, whereas attitude is a disposition within the organism considered without regard to its origins or source of energy. (2) An attitude may be either specific or diffuse in its reference, whereas a sentiment is centered around some definite object. One could speak, for example, of an "antisocial attitude," but scarcely of an antisocial sentiment. (3) Sentiments are conceived as more lasting and hierarchical than attitudes need be. The transitory *Aufgabe* is an attitude as truly as is an enduring philosophy of life, but it is not a sentiment. (4) Sentiment is conscious and benign as contrasted to *complex,* which is a morbid and repressed sentiment. Attitude could cover both wholesomely conscious and morbidly repressed dispositions.

The concept of sentiment received its first systematic use in the hands of Shand (1896) whose theory that human character is composed of sentiments was adopted and elaborated by McDougall. Not only does sentiment seem to these authors to provide a proper unit for characterizing man's social attachments, but also to provide a proper unit for the theory of personality. To McDougall the unity of personality is guaranteed by the possession of the crowning *self-regarding sentiment* at the top of the hierarchical structure.

Certain recent authors likewise have expressed strong preference for the concept of sentiment and strong dislike for the attitude unit. To Murray and Morgan (1945), for example, "attitude" seems too superficial, too motor, to represent the basic organization of needs and cathexes that is, to their minds, the core of "sentiment."

Final word. The four types of units we have discussed — instinct, habit, attitude, sentiment — all claim that dynamic dispositions of some type underlie social behavior. They all have the weakness of assuming a fixity of disposition in the person, and overlook the flexibility of behavior when environmental situations alter. This weakness accounts for the strong attack leveled upon all conative units by field theorists and phenomenologists. In recent years the pendulum has swung away from emphasis on the static unit to a field or situational determinism. Yet individuals in varying situations are not altogether different; the same dispositions tend to be manifest in varying fields. In correcting for the exaggerated fixity implied in these units, it is not necessary to abandon them altogether in favor of momentary situationism, or in favor of an unqualified field theory.

THE BEGINNINGS OF OBJECTIVE METHOD

Experimentation. The year 1879 saw the formal establishment of the first laboratory of psychology. Credit goes to Wundt at Leipzig, although both James at Harvard and Wundt had demonstrational (not research) laboratories as early as 1875. But it took several years longer for social variables to enter the laboratory. The first experimental problem — indeed the only problem for the first three decades of experimental research — was formulated as follows: *What change in an individual's normal solitary performance occurs when other people are present?*

The first laboratory answer to this question came from Triplett (1897). While examining the official records of bicycle races Triplett noted that a rider's maximum speed was approximately 20 percent greater when he was paced by a visible multicycle. Desiring to learn more about the matter, he set up an experiment with children in the age range of ten to twelve, giving them the task of winding fish reels. Alternating situations *alone* and *together,* he found that when working together 20 of his 40 subjects excelled their own solitary record, while 10 did less work, apparently because they were overstimulated by the desire to win, and 10 were essentially unaffected. All in all, he concluded that the group situation must normally be thought of as producing greater output of energy and achievement. His explanation is of interest:

> The bodily presence of another contestant participating simultaneously in the race serves to liberate latent energy not ordinarily available. . . . The sight of the movements of the pacemakers or leading competitors, and the idea of higher speed furnished by this or other means, are probably in themselves dynamogenic factors of some consequence. (p. 533)

In his experimental design, as well as in his explanation, Triplett fails to distinguish two causal factors: emotional competitiveness on the one hand, and simple dynamogenic effect resulting from the sights and sounds of coworkers, on the other. We have already marked how dominant the ideo-motor theory was in late nineteenth-century psychology. Automatic and unconscious movements resulting from implanted ideas or from mere sensory or affective stimulation, were objects of great interest. Féré (1900), a physician in Charcot's clinic, Binet (1900), and others, had great faith in the explanatory value of the concept. Automatisms, mind-reading, suggestion, imitation, hypnotic phenomena were all traced to dynamogenic effects. It was natural to explain group influence in this way.

The same basic problem was an early concern of pedagogics. A. Mayer (1903) studied both the quantity and quality of school children's homework and classwork with respect to memorization, composition, arithmetic, and other tasks. On the whole, his results clearly favored group over solitary work, although he found, as Triplett did, that conditions of overkeen competition in congregate groups affected adversely the quality of the production.

It was a full generation later that the important distinction was drawn between emulation or competition on the one hand, and simple "social facilitation" (whether or not explained in terms of dynamogenesis) on the other. F. H. Allport (1924), who drew this distinction, also pointed to the existence of two kinds of small groups: the *co-acting* and the *face-to-face*. In the former one finds the simplest psychology operating: people working side by side receive their social stimulation almost wholly from "contributory" social stimuli and thus most clearly display the effects of simon-pure "social facilitation." In face-to-face groups, on the other hand, entirely new problems arise: conversation, interaction, circular social behavior, "group thinking," and all manner of direct (not merely contributory) social effects. Accepting this distinction, we may say that early work dealt entirely with co-acting groups. It took a longer time for face-to-face groups to become a central focus of interest. Bechterew and de Lange (1924), G. Watson (1928), and, especially, Kurt Lewin (see collections of his studies, 1948, 1951) gave incentive to research in face-to-face groups.

In 1913 Moede began a monumental program of research on co-acting groups. He decided that a systematic attempt should be made to introduce the social variable into all, or nearly all, standard psychological experiments. If thresholds for the intensity of audible sound could be determined for the individual working alone, why not for two, four, six, or any number of subjects working in the presence of one another? Also, why should not the available instrumental devices for recording involuntary muscular movements be used in the study of imitation? The subject's involuntary movements could, for example, be measured while he was watching similar movements of the experimenter's arm. Also associative processes, shifts in the fixation of attention, learning and forgetting — all well-known phenomena in experimental psychology — could be (and were) studied, using groups of Leipzig school and college students as subjects. Moede published his results in *Experimentelle Massenpsychologie* (1920), but his pioneer conception had less influence than it should, partly because he failed to generalize his findings into systematic theory, and partly because his stimulating book was not translated into English.

Moede's experimental program was, however, known to Münsterberg at Harvard. It was he who, in 1915, encouraged F. H. Allport to undertake work in this area. The latter confirmed the principal findings of Moede and his predecessors and formulated a series of generalizations (1920). Even when the effects of competitiveness are minimized the mere presence of co-workers seems ordinarily to increase the quantity of output (*social increment*); it tends, however, to make judgments conservative, to make associations less personal, and to diminish the quality of reasoning and other complex mental operations (*social subvaluent*). This work was systematically extended by Whittemore (1924) who added the variable of competitiveness, and found in general that it enhanced all these social effects. Unlike Triplett, these authors explained their results in terms of "contributory social stimulation" — a summation of cues and innervations resulting from previous conditioning and leading to the excess output over solitary production. When the variable of competition is added, emotional reinforcement from within summates with the contributory cues from without.

This type of work, with many elaborations, was carried on by subsequent investigators. Travis, for example, found that some of the effects were reversed in the case of stutterers, whose social shyness seemed to be an inhibiting variable (1925). Dashiell summarized all this work and added his own research on the *imagined* as compared with the *actual* presence of others in his review of 1935. Among early experiments on group influence was that of H. T. Moore, who gauged the relative effect of expert opinion and majority opinion upon moral and aesthetic judgments (1921). Moore's research initiated a new method. Instead of having "experts" or a "majority" physically present, he merely reported their judgments to his subjects. In this way it was discovered that the effects of social prestige can be measured simply through instructions that induce an attitude. Countless subsequent investigations have taken advantage of this fact, and have measured social effects through the artificial arousal of preexisting attitude.

Russian psychology, also employing the experimental method, soon after the Bolshevik revolution turned its attention to problems of collective vs. individual behavior. Bechterew and de Lange (1924), using face-to-face (not merely co-acting) groups, report on the effects of discussion in modifying previous individual judgments. Collective thinking, they concluded, is not less efficient than private thinking; often it is superior in terms of the accuracy of judgments. Russian investigations centered attention likewise on the relative merits of competitive and co-operative social situations. As summarized by Boldyreff (1937), their studies established in general the superiority of cooperative arrangements. Recently, however, Bauer (1952) has shown that Soviet psychology has shifted from its earlier emphasis upon the importance of environmental, situational, and co-operative factors in human behavior, to an emphasis upon the duties and obligations of the individual. The superiority of free discussion, co-operation, and teamwork, established by early experiments, seems today less congenial to Communist Party ideology.

In America the narrower formulation of the experimental question gradually gave way. Group influence upon the mental processes of the individual was no longer the sole focus of experimental interest. Lewin, Lippitt, and White (1939) introduced the conception of *social climate* or *group atmosphere.* They formed experimental groups under various types of leadership to approximate authoritarian, democratic, and laissez-faire social atmospheres. Following this historic experiment, interest in group dynamics, group structure, group decision, group cohesion, spread rapidly into laboratory, community, and industrial research. A broader horizon for experimental social psychology opened.

Controlled observation. While the experimental method approaches the ideal design for social research, it is not adapted to all of the problems with which social psychology must wrestle. No doubt most investigators, from the earliest days, would have insisted that their theories and dicta were based on the solid foundation of "facts." Comte, Spencer, Tarde, Durkheim — all theorists — were interpreting social data as they *observed* them. What they lacked was a method for checking and extending their unaided observation.

Perhaps the earliest attempt to improve private observation was through the adoption of questionnaire procedures. Boring (1950) reports uses of this method as early as 1869 in the Berlin schools. Galton (1883) leaned upon questionnaires in his famous study of imagery. Its use on a vast scale commenced with G. Stanley Hall (1891), chiefly in his studies of child psychology. Soon Starbuck (1899) adapted the method for investigations of the religious beliefs and practices of adults. James's work of genius, *Varieties of Religious Experience* (1902), was based in part upon Starbuck's questionnaire data. In the course of time the need for refinements and safeguards brought special methodological researches in matters of wording, interviewing, coding of responses, and scaling. Originally a loose, naïve instrument, the questionnaire has gradually been forged into a tool of considerable precision.

Gradually, too, statistical aids, invented by Quetelet, Galton, Pearson, Fisher, and others came to serve as safeguards for research in social psychology. Devoting himself specifically to the problem of attitude measurement, Thurstone (1927) first introduced statistical sophistication into this important area of investigation. Sociometry, a quasi-quantitative technique invented by Moreno (1934), enabled social psychologists to assess the personal attractions and rejections among members of a group. Following Gallup's first signal success with public opinion polling in the presidental elections of 1936, an elaborate special branch of sampling and opinion research rapidly developed.

Since most of these signal strides in method are of recent date, they do not form a part of our historical story. The fact is that empiricism and positivism did not enter social psychology to any appreciable extent until the decade of the 1920's. The ideals of objectivity and precision then rapidly assumed a dominant position. By 1931 Murphy and Murphy, in the first edition of their *Experimental Social Psychology,* were able to list over eight hundred relevant studies, and in a revised edition (1937) expanded the entries by several hundred additional titles. Hornell Hart (1949) has demonstrated that in the 1930's the upswing of published research began to conform to an exponential curve of increasing output. No tapering off of this accelerating rate of expansion is yet in sight.

TEXTBOOKS IN SOCIAL PSYCHOLOGY

One may learn something concerning the historical background of contemporary social psychology by surveying its textbooks. The accompanying list includes most, if not all, of the textbooks published in English through the year 1952. Many borderline titles are omitted, such as Dewey's *Human Nature and Conduct: an Introduction to Social Psychology* (1922), and Znaniecki's *The Laws of Social Psychology* (1925), on the grounds that such books are less comprehensive than the ordinary textbook. Such exclusions, to be sure, are debatable and to some degree arbitrary.

1908	E. A. Ross,	*Social Psychology: An Outline and Source Book.* New York: Macmillan.
1908	W. McDougall,	*Introduction to Social Psychology.* London: Methuen & Co.

1917 C. A. Ellwood, *An Introduction to Social Psychology.* New York: D. Appleton & Co.
1918 E. S. Bogardus, *Essentials of Social Psychology.* Los Angeles: University of Southern California Press.
1922 J. M. Williams, *Principles of Social Psychology.* New York: Knopf.
1923 R. H. Gault, *Social Psychology: the Basis of Behavior Called Social.* New York: Holt.
1924 F. H. Allport, *Social Psychology.* Boston: Houghton Mifflin.
1924 E. S. Bogardus, *Fundamentals of Social Psychology.* New York: Century Co.
1925 K. Dunlap, *Social Psychology.* Baltimore: Williams & Wilkins.
1925 C. A. Ellwood, *The Psychology of Human Society.* New York: Appleton.
1925 R. H. Thouless, *Social Psychology.* London: Clive.
1925 L. L. Bernard, *Introduction to Social Psychology.* New York: Holt.
1927 K. Young, *Source Book for Social Psychology.* New York: Knopf.
1928 R. Mukerjee and N. N. Sengupta, *Introduction to Social Psychology.* London: Heath.
1929 B. C. Ewer, *Social Psychology.* New York: Macmillan.
1929 J. R. Kantor, *An Outline of Social Psychology.* Chicago: Follett Publishing Co.
1929 C. Murchison, *Social Psychology: the Psychology of Political Domination.* Worcester, Mass.: Clark University Press.
1930 K. Young, *Social Psychology.* New York: Crofts.
1931 J. K. Folsom, *Social Psychology.* New York: Harper & Bros.
1931 E. T. Krueger and W. C. Reckless, *Social Psychology.* New York: Longmans.
1931 G. and Lois B. Murphy, *Experimental Social Psychology.* New York: Harper & Bros. (Revised Edition, G. & Lois B. Murphy & T. M. Newcomb. 1937)
1934 L. G. Brown, *Social Psychology: the Natural History of Human Nature.* New York: McGraw-Hill.
1934 A. Meyerson, *Social Psychology.* New York: Prentice-Hall.
1935 C. Murchison, (Ed.) *Handbook of Social Psychology.* Worcester, Mass.: Clark University Press.
1935 J. J. Smith, *Social Psychology.* Boston: Bruce Humphries.
1936 J. F. Brown, *Psychology and the Social Order.* New York: McGraw-Hill.
1936 E. Freeman, *Social Psychology.* New York: Holt.
1936 H. Gurnee, *Elements of Social Psychology.* New York: Farrar & Rinehart.
1936 R. T. La Piere and P. R. Farnsworth, *Social Psychology.* New York: McGraw-Hill.
1938 D. Katz and R. L. Schanck, *Social Psychology.* New York: Wiley.
1938 J. M. Reinhardt, *Social Psychology.* Philadelphia: Lippincott.
1938 R. T. La Piere, *Collective Behavior.* New York: McGraw-Hill.
1940 C. Bird, *Social Psychology.* New York: Appleton Century.
1940 O. Klineberg, *Social Psychology.* New York: Holt.
1941 S. H. Britt, *Social Psychology of Modern Life.* New York: Farrar & Rinehart.
1942 M. H. Krout, *Introduction to Social Psychology.* New York: Harper & Bros.
1947 T. M. Newcomb and E. L. Hartley, *Readings in Social Psychology.* New York: Holt. (Revised Edition: G. E. Swanson, T. M. Newcomb, and E. L. Hartley. 1952)
1948 D. Krech and R. S. Crutchfield, *Theory and Problems of Social Psychology.* New York: McGraw-Hill.
1948 M. Sherif, *An Outline of Social Psychology.* New York: Harper & Bros.
1948 M. F. Vaughan, *Social Psychology.* New York: Odyssey Press.
1949 N. S. Phadke, *An Outline of Social Psychology.* Kolhapur: College Press.
1949 A. L. Beeley, *Outlines of Social Psychology.* Utah: University of Utah Press.
1949 A. R. Lindesmith and A. L. Strauss, *Social Psychology.* New York: Dryden Press.
1950 T. M. Newcomb, *Social Psychology.* New York: Dryden Press.
1950 S. S. Sargent, *Social Psychology.* New York: Ronald Press.
1950 O. Friedman, *Introduction to Social Psychology.* London: Sylvan Press.
1951 E. L. Queener, *Introduction to Social Psychology.* New York: William Sloane.
1952 S. E. Asch, *Social Psychology.* New York: Prentice-Hall.
1952 L. W. Doob, *Social Psychology.* New York: Holt.
1952 R. E. L. Faris, *Social Psychology.* New York: Ronald Press.
1952 E. L. Hartley and R. E. Hartley, *Fundamentals of Social Psychology.* New York: Knopf.
1952 W. J. H. Sprott, *Social Psychology.* London: Methuen & Co.

The first fact we note in the listing is a gradual acceleration of output excepting during World War II. The distribution in five-year intervals is as follows:

1908-1912	2
1913-1917	1
1918-1922	2
1923-1927	8
1928-1932	8
1933-1937	8
1938-1942	7
1943-1947	1
1948-1952	15
Total	52

Textbooks may vary in content between two extremes. At one pole are the handbooks, devoted largely to experimental and empirical reporting. These works do not presume to integrate all the research surveyed into coherent theory. Examples are the handbooks edited by Murchison (1935) and Lindzey (1954), and the *Experimental Social Psychology* of Murphy, Murphy, and Newcomb (1937). At the other extreme we have textbooks almost totally devoid of reference to specific investigations. Their authors are interested in interpretation and in broad generalizations. When data are not available, the authors of such texts speculate freely about what may be the laws of social psychology. And, of course, there are all sorts of combinations wherein factual reporting is interspersed with interpretation, generalization, and speculation. In an article entitled "On the making of textbooks in social psychology," L. L. Bernard (1931) defended the latter method. This author argued that single *ad hoc* studies are often as misleading as speculation. He points out further that many problems in social psychology are not accessible to experimental study, and that others, while in principle accessible, have not yet been studied experimentally. He likewise raises the question whether experimental validation is necessarily the one heaven-sent proof of assertions in social psychology:

> If one accepts the narrower alternative of putting into a text only those things which conform to one special criterion of validation — say the experimental method — he lays himself open to serious difficulties in addition to the limited character of his product. Who shall decide what is the proper criterion of validation? The experimentalists may naïvely answer that the experimental method is its own justification, because its results can be duplicated by any other worker under the same conditions. I am inclined to think that those who would urge a Methodist, a Catholic, or a Freudian criterion of validation could make the same argument. If one will put himself in the same attitude of mind as one who urges one or the other of these criteria, he will unquestionably be convinced by the same evidence. (p. 72)

Bernard here warns us that ultimately deep epistemological decisions underlie the building of social psychology as a science. To assume that *ad hoc* experiments are themselves the final arbiter of truth is, as he maintains, naïve. No experiment interprets itself; no fact speaks for itself. Theories that transcend the specific instance are necessary, and theory building in social psychology is still in its infancy.

Textbooks vary not only in their addiction to findings or to broad interpretation, but also in respect to their systematic defense of a single point of view. Some authors are happy with a loose eclecticism; they worry not at all about the coherence between findings, or even between theories. Others, by contrast, are intent upon constructing a framework that will bind facts and theories into a superordinate and logical whole. Among the textbooks that can clearly be called systematic are those of McDougall (1908), F. H. Allport (1924), Brown (1936), Sherif (1948), Krech and Crutchfield (1948), Asch (1952). Most authors, no doubt, would claim that they are presenting a systematic point of view. But often their product is merely a matter of weighted preference for certain concepts. Genuinely systematic texts are those that tie the data of social psychology to a coherent psychological or sociological frame of thought broader than the data of social psychology itself. Thus McDougall's social psychology is wholly consistent with his over-all hormic approach to all psychological problems; F. H. Allport's text is consistent with the behaviorism prevailing in 1924. J. F. Brown favors field theory; Sherif, cultural determinism; Krech and Crutchfield a phenomenological approach; Asch, Gestalt psychology.

A final and chief cause of the great diversity in textbooks lies in the sociological-psychological bifurcation that marks the entire history of social psychology. In the accompanying list, by rough estimate, two-thirds of the texts are written by psychologists and one-third by sociologists. Many, therefore, reflect a strong psychological bias, and many a sociological. What is the difference?

Earlier in the present essay we ventured to define social psychology as the attempt to understand and explain how the thought, feeling, and behavior of individuals are influenced by the actual, imagined, or implied presence of other human beings. This definition has an obviously psychological cast. By contrast, we may refer to the following sociological type of definition. Ellwood (1925) writes:

> Social psychology is the study of social interaction. It is based upon the psychology of group life. It begins with an interpretation of group-made types of human reactions, of communication, and instinctive and habitual actions. (p. 16)

It would be invidious to designate each and all of the texts in the accompanying list as either psychological or sociological in approach. Some authors would be sure to protest against the typing. But there would be little disagreement in designating as psychologistic the texts by McDougall (1908) and by F. H. Allport (1924). Nor would there be much opposition to labeling as sociologistic the texts by Ellwood (1917) or Lindesmith and Strauss (1949). Combination points of view are, of course, possible, and are increasing in frequency. Newcomb (1950) might be said to hold with equal tenacity to a social and to a psychological frame of reference.

FINAL WORD

The diversity of textbooks is, of course, a natural reflection of the diversity that has marked the whole history of social psychology. Writers have differed in their definitions of the field, in their selection of problems for study, and above all in their conceptual tools and presuppositions. The student of social psychology has no choice but to learn many maps, realizing that the master chart is not yet available. Many maps, however, are preferable to no maps. Investigators who prefer to start from scratch do so at their own peril. When they attempt to put their myriad empirical findings together they may find that they are entering blind alleys that have trapped many in the past; or they may find that their "discoveries" are, historically considered, mere truisms. But the investigator who is aware of the history of his subject is able to strike out with assurance. He will be able to distinguish new garbs from old, progress from platitude, and will thus help build a truly cumulative science.

REFERENCES

ABERLE, D. F., COHEN, A. K., DAVIS, A. K., LEVY, M. J., & SUTTON, F. X. The functional prerequisites of a society. *Ethics,* 1950, *60,* 100-111.

ACH, N. *Ueber die Willenstätigkeit und das Denken.* Göttingen: Vanderhoeck and Ruprecht, 1905.

ADLER, A. A study of organ inferiority and its psychic compensations. Transl. *Nerv. ment. Dis. Monogr. Ser.,* 1917, *24.*

ALLPORT, F. H. The influence of the group upon association and thought. *J. exp. Psychol.,* 1920, *3,* 159-182.

ALLPORT, F. H. *Social psychology.* Boston: Houghton-Mifflin, 1924.

ALLPORT, F. H. *Institutional behavior.* Chapel Hill: Univ. of North Carolina Press, 1933.

ALLPORT, F. H. An event system theory of collective action: with illustrations from economic and political phenomena and the production of war. *J. soc. Psychol., SPSSI Bull.,* 1940, *11,* 417-445.

ALLPORT, G. W. Attitudes. In C. M. Murchison (Ed.), *Handbook of Social Psychology.* Worcester, Mass.: Clark Univ. Press, 1935. Pp. 798-844.

ALLPORT, G. W. *Personality: a psychological interpretation.* New York: Henry Holt, 1937.

ALLPORT, G. W. Dewey's individual and social psychology. In P. A. Schilpp (Ed.), *The philosophy of John Dewey.* Evanston, Ill.: Northwestern Univ. Press, 1939. Pp. 265-290.

ALLPORT, G. W. The ego in contemporary psychology. *Psychol. Rev.,* 1943, *50,* 451-478.

ALLPORT, G. W. Effect: a secondary principle of learning. *Psychol. Rev.,* 1946, *53,* 335-347.

ALLPORT, G. W. A psychological approach to the study of love and hate. In P. A. Sorokin (Ed.), *Explorations in altruistic love and behavior: a symposium.* Boston: Beacon Press, 1950. Pp. 145-164.

ALLPORT, G. W. The trend in current theories of motivation. *Amer. J. Orthopsychiat.,* 1953, *23,* 107-119.

ARISTOTLE. *Politics.* Transl. by B. Jowett. Oxford: Clarendon Press, 1885.

ARNOLD, T. *Folklore of capitalism.* New Haven: Yale Univ. Press, 1937.

ASCH, S. E. The doctrine of suggestion, prestige, and imitation in social psychology. *Psychol. Rev.,* 1948, *55,* 250-276.

ASCH, S. E. *Social psychology.* New York: Prentice-Hall, 1952.

ASHLEY-MONTAGU, M. F. *On being human.* New York: H. Schuman, 1950.

BACON, F. *Novum organum.* (First published in 1620.) Transl. London & New York: The Colonial Press, 1900. Sections 38-68.

BAGEHOT, W. *Physics and politics.* New York: D. Appleton, 1875.

BAKER, H. *The dignity of man.* Cambridge: Harvard Univ. Press, 1947.

BALDWIN, J. M. *Mental development in the child and in the race.* New York: Macmillan, 1895.

BALDWIN, J. M. *Social and ethical interpretations in mental development.* New York: Macmillan, 1897.

Barnes, H. E. *Historical sociology, its origins and development.* New York: Philosophical Library, 1948.

Barnes, H. E. (Ed.) *An introduction to the history of sociology.* Chicago: Univ. of Chicago Press, 1948.

Bartlett, F. C. *Psychology and primitive culture.* Cambridge, Eng.: Cambridge Univ. Press, 1923.

Bartlett, F. C. *Remembering.* Cambridge, Eng.: Cambridge Univ. Press, 1932.

Bastian, A. *Der Mench in der Geschichte: zur Begründung einer psychologischen Weltanschauung.* 3 vols. Leipzig: O. Wigand, 1860.

Bauer, R. A. *The new man in Soviet psychology.* Cambridge: Harvard Univ. Press, 1952.

Bechterew, V. M. *General principles of human reflexology.* (Transl. from 4th Russian Ed., 1928.) New York: International Publishers, 1932.

Bechterew, V. M., & de Lange, M. Die Erbegnisse des Experiments auf dem Gebiete der kollektiven Reflexologie. *Z. angew. Psychol.,* 1924, *24,* 305-344.

Becker, H. Some forms of sympathy; a phenomenological analysis. *J. abnorm. soc. Psychol.,* 1931, *26,* 58-68.

Becker, H., & Barnes, H. E. *Social thought from lore to science* (2nd Ed.). Washington: Hansen Press, 1952.

Bentham, J. *An introduction to the principles of morals and legislation.* (First published 1789.) Oxford: Clarendon Press, 1879.

Bernard, L. L. *Instinct: a study in social psychology.* New York: Henry Holt, 1926.

Bernard, L. L. On the making of textbooks in social psychology. *J. educ. Sociol.,* 1931-32, *5,* 67-81.

Bernard, L. L. Social psychology, in *Encyclopedia of social science.* Vol. 14. New York: Macmillan, 1934. Pp. 151-157.

Bernheim, H. *De la suggestion dans l'état hypnotique et dans l'état de veille.* Paris: O. Doin, 1884.

Binet, A. *La suggestibilité.* Paris: Schleicher, 1900.

Blanton, S., & Blanton, Margaret. *Child guidance.* New York: Century Co., 1927.

Bogardus, E. S. *A history of social thought.* Los Angeles: Univ. of Southern California Press, 1922.

Boldyreff, J. W. Recent materials on competition and cooperation in Russia. In M. A. May, G. W. Allport, G. Murphy (Eds.), *Memorandum on research in competition and cooperation.* New York: Social Science Research Council, 1937. (Mimeo.)

Boring, E. G. *A history of experimental psychology.* New York: D. Appleton-Century, 1929. Rev. Ed., Appleton-Century-Crofts, 1950.

Bosanquet, B. *Philosophical theory of the state.* New York: Macmillan, 1899.

Bousquet, G. H. *The work of V Pareto.* Hanover, N. H.: The Sociological Press, 1928.

Braid, J. *Neurypnology.* (Originally published 1843.) Rev. ed., London: G. Redway, 1899.

Brett, G. S. *A history of psychology.* 3 Vols. London: G. Allen, 1912-1921.

Britt, S. H. *Selected readings in social psychology.* New York: Rinehart, 1950.

Brown, T. *Lectures on the philosophy of the human mind.* 4 Vols. Edinburgh: J. Ballantyne for W. Tait and C. Tait, 1820.

Cantril, H. *The invasion from Mars.* Princeton: Princeton Univ. Press, 1940.

Cantril, H. *The psychology of social movements.* New York: Wiley, 1941.

Catlin, G. *The story of the political philosophers.* Rev. ed., New York: Tudor, 1947.

Charcot, J. M. *Oeuvres complètes.* 9 Vols. Paris: Bureaux du Progrès médical, 1888-1894.

Chase, S. *Roads to agreement.* New York: Harper, 1951.

Clarke, H. M. Conscious attitudes. *Amer. J. Psychol.,* 1911, *32,* 214-249.

Comte, A. *The positive philosophy.* Vol. I. (Original French Ed., 1830.) Transl. London: Trubner & Co., 1853.

Comte, A. *The positive polity.* (Original French Ed., Vol. 1, 1852; Vol. 4, 1854.) Transl., London: Longmans, Green, 1875.

Condorcet (Caritat, M. J. A. N.). *A sketch of the intellectual progress of mankind.* Transl., London: J. Johnson, 1795.

Cooley, C. H. *Human nature and the social order.* New York: Scribners, 1902.

Cottrell, L. S., & Gallagher, Ruth. Developments in social psychology, 1930-1940. *Sociometry Monogr.,* No. 1. New York: Beacon House, 1941.

Curti, M. *The growth of American thought.* New York: Harper, 1943.

Darwin, C. *On the origin of species.* London: J. Murray, 1859.

Dashiell, J. F. Experimental studies of the influence of social situations on the behavior of individual human adults. In C. C. Murchison (Ed.), *Handbook of social psychology.* Worcester, Mass.: Clark Univ. Press, 1935. Pp. 1097-1158.

De Grange, McQ. *The science of individuality.* Lyon: L. Bascou, 1923.

De Grange, McQ. *The curve of societal development.* Hanover, N. H.: Sociological Press, 1930.

De Grange, McQ. The method of Auguste Comte. In S. Rice (Ed.), *Methods in social science.* Chicago: Univ. of Chicago Press, 1931. Pp. 19-58.

Delaisi, F. *Political myth and economic reality.* New York: Viking Press, 1927.

Dewey, J. *Psychology as philosophic method.* Berkeley: Univ. Chronicle, 1899.

Dewey, J. The need for social psychology. *Psychol. Rev.,* 1917, *24,* 266-277.

Dewey, J. *Human nature and conduct: an introduction to social psychology.* New York: Henry Holt, 1922.

Dollard, J., & Miller, N. E. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.

Doob, L. W. *Propaganda: its psychology and techniques.* New York: Henry Holt, 1935.

Droba, D. D. The nature of attitude. *J. soc. Psychol.,* 1933, *4,* 444-463.

Dunlap, K. Are there any instincts? *J. abnorm. Psychol.,* 1919, *14,* 307-311.

Dunlap, K. *Social psychology.* Baltimore: Williams & Wilkins, 1925.

Durkheim, E. Réprésentations individuelles et réprésentations collectives. *Révue de Métaphysique,* 1898, *6,* 274-302.

Durkheim, E. *Le suicide.* Paris: F. Alcan, 1897. Transl., Glencoe, Ill.: Free Press, 1951.

Durkheim, E. *De la division du travail social.* Paris: F. Alcan, 1902. Transl., New York: Macmillan, 1933.

Durkheim, E. *Les formes élémentaires de la vie religieuse.* Paris: F. Alcan, 1912. Transl., New York: Macmillan Co., 1915.

Ellwood, C. A. *The psychology of human society.* New York: Appleton, 1925.

Ellwood, C. A. *A history of social philosophy.* New York: Prentice-Hall, 1938.

Engelmann, G. *Political philosophy from Plato to Jeremy Bentham.* Transl., New York: Harper, 1927.

Espinas, A. *Des sociétés animales.* Paris: G. Baillière. 1877.

Faraday, M. Experimental investigation of table-moving. *Athenaeum,* 1853, *1340,* 801-803.

Faris, E. The concept of social attitudes. *J. appl. Sociol.,* 1925, *9,* 404-409.

Faris, E. L. The beginnings of social psychology. *Amer. J. Sociol.,* 1945, *50,* 422-428.

Féré, C. Note sur la physiologie de l'attention. *Rev. Phil.,* 1890, *30,* 393-405.

Féré, C. *Sensation et mouvement.* Paris: F. Alcan, 1900.

Fouillée, A. *Morale des idées-forces.* (2nd Ed.). Paris: F. Alcan, 1908.

Freud, S. *Wit and its relation to the unconscious.* Transl., New York: Moffat, Yard, 1916.

Freud, S. *Totem and taboo.* Transl., New York: Moffat, Yard, 1918.

Freud, S. *Group psychology and the analysis of the ego.* Transl., London: International Psycho-analytical Press, 1922.

Freud, S. *Beyond the pleasure principle.* Transl., London: International Psycho-analytical Press, 1922.

Galton, F. *Inquiries into human faculty and its development.* London: Macmillan, 1883.

Gehlke, C. E. *Emile Durkheim's contributions to sociological theory.* New York: Columbia Univ. Press, 1915.

Giddings, F. H. *The principles of sociology.* New York: Macmillan, 1896.

Godwin, W. *Enquiry concerning political justice.* London: G. G. J. Robinson & J. Robinson, 1793.

Goldenweiser, A. *History, psychology, and culture.* New York: Knopf, 1933.

Green, T. H. Lectures on the principles of political obligation. In *Collected Works.* Vol. 2. London: Longmans, Green, 1900.

Greenwood, E. *Experimental sociology.* New York: Kings Crown Press, 1945.

Haddon, A. C. *History of anthropology.* London: Watts, 1910; Rev. Ed., 1949.

Hall, G. S. The contents of children's minds on entering school. *Ped. Sem.,* 1891, *1,* 139-173.

Hamilton, W. *Lectures on metaphysics.* (H. Maurel & J. Veitch, Eds.) Boston: Gould & Lincoln, 1859-60.

Hart, H. The pre-war upsurge in social science. *Amer. sociol. Rev.,* 1949, *14,* 599-607.

Hayakawa, S. I. *Language in action.* New York: Harcourt, Brace, 1941.

Hegel, G. W .F. *Phänomenologie des Geistes.* (First published 1807.) Transl., *The phenomenology of mind.* London: G. Allen & Unwin, 1910.

Herbart, J. F. Ueber einige Beziehungen zwischen Psychologie und Staatswissenschaft (1821). In Vol. 5 (1890) of *Sämtliche Werke.* Langensalza: Hermann Beyer, 19 vols., 1887-1912.

Hobbes, T. *Leviathan.* Reprint of 1st (1651) Ed., Cambridge, Eng.: Univ. Press, 1904.

Hocking, W. E. *Man and the state.* New Haven: Yale Univ. Press, 1926.

Hollingworth, H. L. *Psychology of the functional neuroses.* New York: D. Appleton, 1920.

Holt, E. B. *The Freudian wish and its place in ethics.* New York: Holt, 1915.

Holt, E. B. *Animal drive and the learning process.* New York: Holt, 1931.

Hoppe, F. Das Anspruchsniveau. In Untersuchungen zur Handlungs-und Affektpsychologie, IX. Erfolg und Misserfolg. *Psycholog. Forsch.,* 1930, *14,* 1-62.

Horney, K. *New ways in psychoanalysis.* New York: Norton, 1939.

House, F. N. *Development of sociology.* New York: McGraw-Hill, 1936.

Hull, C. L. *Hypnosis and suggestibility.* New York: D. Appleton Century, 1934.

Humphrey, G. Imitation and the conditioned reflex. *Ped. Sem.,* 1921, *28,* 1-21.

Jaensch, E. A. *Der Gegentypus.* Leipzig: Barth & Co., 1938.

Jahoda, Marie, Deutsch, M., & Cook, S. W. *Research methods in social relations.* 2 vols. New York: Dryden Press, 1951.

James, W. *The varieties of religious experience.* New York: Longmans, Green, 1902. (The quotation is from the Modern Library edition.)

James, W. *Principles of psychology.* 2 vols. New York: Holt, 1890.

Jones, E. Rationalization in everyday life. *J. abnorm. Psychol.,* 1908, *3,* 161-169.

Judd, C. H. *The psychology of social institutions.* New York: Macmillan, 1926.

Jung, C. G. *Collected papers on analytical psychology.* (2nd Ed.) Transl., London: Baillière, Tindall & Cox, 1922.

Karpf, F. B. *American social psychology.* New York: McGraw-Hill, 1932.

Karpf, F. B. American social psychology — 1951. *Amer. J. Sociol.,* 1952, *2,* 187-193.

Katona, G. *Psychological analysis of economic behavior.* New York: McGraw-Hill, 1951.

Katz, D. (Ed.) *Handbuch der Psychologie.* Transl. Basel: Schwabe, 1951.

Kempf, E. J. *The autonomic functions and the personality.* New York: Nervous & Mental Disease Pub. Co., 1918.

Kluckhohn, C. *Navaho witchcraft.* Cambridge: Peabody Museum Papers, 1944, *22,* No. 2.

Kluckhohn, C., & Murray, H. A. (Eds.) *Personality in nature, society, and culture.* New York: A. Knopf, 1948; Rev. Ed., 1952.

Koffka, K. *Zur Analyse der Vorstellungen und ihren Gesetze.* Leipzig: Quelle & Meyer, 1912.

Köhler, W. *The mentality of apes.* Transl., New York: Harcourt, Brace, 1927.

Krech, D., & Crutchfield, R. S. *Theory and problems of social psychology.* New York: McGraw-Hill, 1948.

Kroeber, A. L., & Kluckhohn, C. *Culture: a critical review of concepts and definitions.* Cambridge: Peabody Museum Papers, 1952, *47,* No. 1.

Kropotkin, P. A. *Mutual aid, a factor of evolution.* New York: McClure, Phillips, 1902.

Lange, L. Neue Experimente über den Vorgang der einfachen Reaktion auf Sinneseindrücke. *Phil. Stud.,* 1888, *4,* 472-510.

Lange, N. Beiträge sur Theorie der sinnlichen Aufmerksamkeit und der activen Apperception. *Phil. Stud.,* 1888, *4,* 390-422.

Lazarus, M., & Steinthal, H. *Zeitschrift f. Völkerpsychologie und Sprachwissenschaft.* Vols. 1-20. Berlin: F. Dümmler, 1860-1890.

Le Bon, G. *Psychologie des foules.* Paris: F. Olean, 1895. Transl., *The Crowd,* London: T. Fisher Unwin, 1896. (Quotations are from the 12th English impression, 1920.)

Lecky, P. *Self-consistency: a theory of personality.* New York: Island Press, 1945.

Le Dantec, F. *L'égoisme: seul base de toute société.* Paris: Flammarion, 1918.

Lee, Vernon (Paget, Violet). *Vital lies.* London & New York: J. Lane, 1912.

Leighton, A. H. *The governing of men.* Princeton: Princeton Univ. Press, 1945.

Lévy-Bruhl, L. *La mentalité primitive.* Paris: F. Alcan, 1922. Transl., *How natives think.* London: S. Allen & Unwin, 1926.

Lévy-Bruhl, L. *The philosophy of Auguste Comte.* Transl., New York: G. P. Putnam's Sons, 1903.

Lewin, K. *Resolving social conflict: selected papers on group dynamics.* New York: Harper, 1948.

Lewin, K. *Field theory in social science.* New York: Harper, 1951.

Lewin, K., Lippitt, R., & White, R. Patterns of aggressive behavior in experimentally created "social climates." *J. soc. Psychol.,* 1939, *10,* 271-299.

Liébeault, A. A. *Du sommeil et des états analogues, considérés surtout au point de vue de l'action de la morale sur le physique.* Paris: V. Masson, 1866.

Lindzey, G. (Ed.) *Handbook of social psychology.* 2 vols. Cambridge: Addison-Wesley, 1954.

Lippmann, W. *Public opinion.* New York: Harcourt, Brace, 1922.

Lipps, T. Das wissen von fremden Ichen. *Psychol. Untersuchungen,* 1907, *1,* 694-722.

Locke, J. *Second treatise on government.* First published 1689. In *Works,* Book II. London: Churchill, Manship, 1714. Chap. 8, Sec. 119.

Lowenthal, L., & Guterman, N. *Prophets of deceit.* New York: Harper, 1949.

Lukasczyk, K. *Der Gegenstand der Sozialpsychologie.* Munich: Ludwig-Maximilians-Universität, 1951. (Mimeo.)

McDougall, W. *Introduction to social psychology.* London: Methuen & Co., 1908.

McDougall, W. *The group mind.* New York: G. P. Putnam's Sons, 1920. Rev. Ed., 1928.

McDougall, W. *Outline of psychology.* New York & Chicago: C. Scribner's, 1923.

McDougall, W. *The energies of men: a study of the fundamentals of dynamic psychology.* London: Methuen, 1932.

McDougall, W. Pareto as a psychologist. *J. soc. Phil.,* 1935, *1,* 36-52.

McDougall, W. Tendencies as indispensable postulates of all psychology. Paris: *Proceedings of the XIth International Congress on Psychology,* 1937, 157-170.

Machiavelli. *The prince.* First published, 1513. Transl., New York: Hendricks House, 1946.

MacLeod, R. The place of phenomenological analysis. In J. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads.* New York: Harper, 1951. Pp. 215-241.

Maier, N. R. F. *Principles of human relations.* New York: Wiley, 1952.

Malinowski, B. *A scientific theory of culture and other essays.* Chapel Hill: Univ. of North Carolina Press, 1944.

Mannheim, K. *Ideology and utopia.* Transl., New York: Harcourt, Brace, 1936.

Martin, E. D. *The behavior of crowds.* New York: Harper, 1920.

Marx, K., & Engels, F. *The German ideology.* (Written in 1846; first portions published in 1903-04.) Transl., New York: International Publishers, 1937.

May, M. A., & Doob, L. W. *Competition and cooperation.* New York: Social Science Research Council, 1937, Bull. No. 25.

Mayer, A. Über Einzel- und Gesamtleistung des Schulkindes. *Arch. ges. Psychol.,* 1903, *1,* 276-416.

Mead, G. H. *Mind, self and society.* (Posthumous; C. M. Morris, Ed.) Chicago: Univ. of Chicago Press, 1934.

Mead, Margaret. *Cooperation and competition among primitive peoples.* New York: McGraw-Hill, 1937.

Meltzer, H. Present status of experimental studies on the relationship of feeling to memory. *Psychol. Rev.,* 1930, *37,* 124-129.

Mesmer, F. A. *Mémoire sur la découverte du magnétisme animal.* First published, 1779. Transl., *Mesmerism.* London: MacDonald, 1948.

MILL, J. S. *Utilitarianism.* Reprint from *Fraser's Magazine,* 1861, *64;* London: Parker, son & Bourn, 1863.

MILLER, N. E., & DOLLARD, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.

MOEDE, W. *Experimentelle Massenpsychologie.* Leipzig: S. Hirzel, 1920.

MOORE, H. T. The comparative influence of majority and expert opinion. *Amer. J. Psychol.,* 1921, *32,* 16-20.

MORENO, J. L. *Who shall survive?* Washington, D.C.: Nervous & Mental Disease Publ. Co., 1934.

MÜLLER, G. E., & PILZECKER, A. Experimentelle Beiträge zur Lehre vom Gedächtniss. *Z. Psychol.,* 1900, Ergbd. 1.

MÜNSTERBERG, H. *Beiträge zur experimentellen Psychologie.* Vol. 1. Freiburg: Morh, 1889.

MURCHISON, C. C. (Ed.) *Handbook of social psychology.* Worcester, Mass.: Clark Univ. Press, 1935.

MURRAY, H. A., MORGAN, CHRISTIANA. A clinical study of sentiments. *Genet. Psychol. Monogr.,* 1945, *32,* 3-149.

MURPHY, G. *Historical introduction to modern psychology.* New York: Harcourt, Brace, 1939; Rev. Ed., 1949.

MURPHY, G., & MURPHY, LOIS B. *Experimental social psychology.* New York: Harper, 1931; Rev. Ed. by G. Murphy, Lois B. Murphy, & T. M. Newcomb, 1937.

MURPHY, LOIS B. *Social behavior and child personality: an exploratory study of some roots of sympathy.* New York: Columbia Univ. Press, 1937.

NIETZSCHE, F. *Der Wille zur Macht.* Book 3, sec. 702. In Nietzsche's complete *Werke,* vol. 16. Leipzig: Alfred Kröner, 1912.

OGBURN, W. F. *Social change with respect to culture and original nature.* New York: B. W. Huebsch, 1922; Rev. Ed., New York: Viking Press, 1950.

OGDEN, C. K. *Bentham's theory of fictions.* London: K. Paul, Trench, Truber, 1932.

PARETO, V. *The mind and society.* 1st Italian Ed., 1916. Transl., New York: Harcourt, Brace, 1935.

PARK, R. E. Experience and race relations. *J. appl. Sociol.,* 1924, *9,* 18-24.

PARK, R. E., & BURGESS, E. W. *Introduction to the science of sociology.* Chicago: Univ. of Chicago Press, 1924.

PARSONS, T. *The structure of social action.* New York: McGraw-Hill, 1937.

PARSONS, T. The rise and decline of economic man. *J. gen. Educ.,* 1949, *4,* 47-53.

PARSONS, T., & SHILS, E. A. (Eds.) *Toward a general theory of action.* Cambridge: Harvard Univ. Press, 1951.

PEAR, T. H. (Ed.) *Psychological factors of peace and war.* London: Hutchinson, 1950.

PERRY, R. B. Is there a social mind? *Amer. J. Sociol.,* 1922a, *27,* 561-572.

PERRY, R. B. Is there a social mind? *Amer. J. Sociol.,* 1922b, *27,* 721-736.

PIAGET, J. *The moral judgment of the child.* Transl., London: Paul, Trench, Truber, 1932.

PLATO. *The republic.* Vol. 3 of *Dialogues of Plato* (5 vols.). Transl. by B. Jowett (2nd Ed.). Oxford: Clarendon Press, 1875.

RIBOT, T. *The psychology of the emotions.* Transl., London: W. Scott, Ltd., 1897.

RICE, S. A. (Ed.) *Methods in social science.* Chicago: Univ. of Chicago Press, 1931.

RIVERS, W. H. R. *Psychology and ethnology.* New York: Harcourt, Brace, 1926.

ROBINSON, J. H. *The mind in the making.* New York: Harper, 1921.

ROGERS, C. R. *Counseling and psychotherapy.* Boston: Houghton Mifflin, 1942.

RÓHEIM, G. *Australian totemism, a psychoanalytic study in anthropology.* London: G. Allen & Unwin, 1925.

ROSS, E. A. *Social psychology.* New York: Macmillan, 1908.

ROYCE, J. *The problem of Christianity.* Vol. II. New York: Macmillan, 1913.

RUSSELL, B. *Power.* New York: Norton, 1938.

RUSSELL, B. *History of western philosophy.* New York: Simon & Schuster, 1945.

SABINE, G. *A history of political theory.* New York: Henry Holt, 1937.

SCHAEFFLE, A. *Bau und Leben des sozialen Körpers.* 4 vols. Tübingen: H. Laupp, 1875-78.

SCHELER, M. *Wesen und Formen der Sympathie.* (2nd Ed.) Bonn: Friedrich Cohen, 1923.

SCHOPENHAUER, A. *Die Welt als Wille und Vorstellung.* Leipzig: F. H. Brockhaus, 1819. Transl., *The world as will and idea.* (6th Ed.) London: Kegan Paul, Trench, Trubner, 1909. Append., Chap. 19.

SHAND, A. Character and the emotions. *Mind,* 1896, *21,* 203-342.

SIDIS, B. *The psychology of suggestion.* New York: D. Appleton, 1898.

SIGHELE, S. *La foule criminelle.* Original Italian Ed., 1891. Transl., Paris: Baillière, 1892.

SIGHELE, S. *Le crime à deux.* Original Italian Ed., 1893. Transl., Lyon: A. Storck, 1893.

SIGHELE, S. *Psychologie des sectes.* Original Italian Ed., 1895. Transl., Paris: V. Giard et E. Brière, 1898.

SIMMEL, G. *The sociology of Georg Simmel.* Transl. & Ed., K. H. Wolff. Glencoe, Ill.: Free Press, 1950.

SMITH, A. *The theory of moral sentiments.* London: A. Miller, 1759.

SMITH, A. *An inquiry into the nature and causes of the wealth of nations.* London: W. Strahan & T. Cadell, 1776.

SOREL, G. *Reflections on violence.* Transl., New York: B. W. Huebsch, 1914.

SOROKIN, P. A. *Contemporary sociological theories.* New York: Harper, 1928.

SOROKIN, P. A. (Ed.) *Explorations in altruistic love and behavior: a symposium.* Boston: Beacon Press, 1950.

Spencer, H. *First principles*. Preface dated 1862. New York: Appleton & Co., 1895.

Spencer, H. *The principles of psychology*. 2 vols. (2nd Ed.) London: Williams & Norgate, 1870-72.

Spencer, H. *The principles of sociology*. Vols. 1-2. First published, 1876. New York: D. Appleton, 1900.

Sprowls, J. W. *Social psychology interpreted*. Baltimore: Williams & Wilkins, 1927.

Starbuck, E. D. *The psychology of religion*. London: W. Scott, 1899.

Stirner, M. (J. K. Schmidt). *Der Einzige und sein Eigenthum*. Leipzig: O. Wigand, 1845. Transl., *The ego and his own*. New York: B. R. Tucker, 1907.

Stratton, G. M. *Experimental psychology and its bearing upon culture*. London & New York: Macmillan, 1903.

Stouffer, S. A. *et al. The American soldier*. 2 vols. Princeton: Princeton Univ. Press, 1949.

Suttie, I. D. *The origins of love and hate*. London: Kegan Paul, Trench, Truber, 1935.

Tarde, G. *The laws of imitation*. Transl., New York: Henry Holt, 1903.

Thomas, W. I., & Zananiecki, F. *The Polish peasant in Europe and America*. 5 vols. Boston: Badger, 1918-20.

Thorndike, E. L. Animal intelligence: an experimental study of the associative process in animals. *Psychol. Rev.*, 1898, Monogr. Suppl. No. 8.

Thorndike, E. L. *Educational psychology*. Vol. 2. *The original nature of man*. New York: Teachers College, Columbia Univ., 1913.

Thurnwald, R. (Ed.) *Forschungen zur Völkerpsychologie und Soziologie*. Vols. 1-14. Leipzig: C. L. Hirschfeld, 1925-35.

Thurstone, L. L. Attitudes can be measured. *Amer. J. Sociol.*, 1927-28, *33*, 529-554.

Titchener, E. B. *Lectures on the elementary psychology of feeling and attention*. New York: Macmillan, 1908.

Titchener, E. B. *Experimental psychology of the thought processes*. New York: Macmillan: 1909.

Titchener, E. B. *A textbook of psychology*. New York: Macmillan, 1910. New Ed., 1916.

Travis, L. E. The influence of the group upon the stutterer's speed in free association. *J. abnorm. soc. Psychol.*, 1925, *20*, 142-146.

Triplett, N. The dynamogenic factors in pacemaking and competition. *Amer. J. Psychol.*, 1897, *9*, 507-533.

Troland, L. T. *Fundamentals of human motivation*. New York: Van Nostrand, 1928.

Trotter, W. *Instincts of the herd in peace and war*. New York: Macmillan, 1916.

Vaughan, C. E. *Studies in the history of political philosophy before and after Rousseau*. 2 vols. Manchester, Eng.: University Press, 1939.

Vico, G. *The new science of Giambattista Vico*. First published 1725. Transl. from 3rd Ed. (1744). Ithaca: Cornell Univ. Press, 1948.

Waelder, R. *Psychological aspects of war and peace*. Geneva: Geneva Research Centre, 1939.

Wallas, G. *Human nature in politics*. London: A. Constable, 1908.

Wallas, G. *The great society*. New York: Macmillan, 1914.

Wallis, W. D. The independence of social psychology. *J. abnorm. soc. Psychol.*, 1925, *20*, 147-150.

Ward, C. H. *Thobbing*. Indianapolis: Bobbs-Merrill, 1926.

Warren, H. C. (Ed.) *Dictionary of psychology*. Boston: Houghton Mifflin, 1934.

Washburn, M. F. *Movement and mental imagery*. Boston: Houghton Mifflin, 1916.

Watson, G. Do groups think more efficiently than individuals? *J. abnorm. soc. Psychol.*, 1928, *23*, 328-336.

Watson, J. B. *Psychology from the standpoint of a behaviorist*. New York: J. B. Lippincott, 1919.

White, L. W. Culturological versus psychological interpretations of human behavior. *Amer. sociol. Rev.*, 1947, *12*, 686-698.

Whittemore, I. C. Influence of competition on performance: an experimental study. *J. abnorm. soc. Psychol.*, 1924, *19*, 236-253.

Wiener, N. *Cybernetics*. New York: John Wiley, 1948.

Wild, J. *Plato's theory of man*. Cambridge: Harvard Univ. Press, 1948.

Williams, W. *What's on the worker's mind*. New York: Scribners, 1920.

Windelband, W. *A history of philosophy*. Transl., (2nd Rev. Ed.), New York: Macmillan, 1935.

Wissler, C. *Man and culture*. New York: Thomas Y. Crowell, 1923.

Woodworth, R. S. *Dynamic psychology*. New York: Columbia Univ. Press, 1918.

Wundt, W. *Beiträge zur Theorie der Sinneswahrnehmung*. Leipzig: C. F. Winter, 1862.

Wundt, W. *Völkerpsychologie: eine Untersuchung der Entwicklungsgesetze von Sprache, Mythus und Sitte*. 10 vols. Leipzig: Engelmann, 1900-1920. Transl. (Vols. 1-3), *Elements of folk psychology*. New York: Macmillan, 1916.

Young, K. (Ed.) *Social attitudes*. New York: Henry Holt, 1931.

Znaniecki, F. *The laws of social psychology*. Chicago: University Press, 1925.

Part 2

Contemporary Systematic Positions

CHAPTER 2

Stimulus-Response Contiguity and Reinforcement Theory in Social Psychology

William W. Lambert*
Cornell University

Two aspects of the varied roles of Stimulus-Response theory in contemporary social psychology will be emphasized here. One general area of discussion will be a summary of S-R contiguity theory and S-R reinforcement theory; the other will be a consideration of the problem of generalizing S-R theory for social psychology. Before these topics are broached specifically, however, some precursors of these theories will be touched on.

For some, the current notions of S-R theorists are direct descendants of various doctrines of the association of ideas; for others the theories are operationally bound games of runways and mazes; and for still others the theories are based on artificial data and are fitting only to the experimental laboratory whence they came. Decision as to whether any of these characterizations is correct will, of course, be an individual matter, but a view of the present range and extensions of S-R theory in social psychology is of quite general interest. Insofar as S-R theory is associationistic, it has a history as long as that of western culture. Insofar as it is couched within a general empiricist philosophy, its history is largely that of the British empiricist school. Insofar as it is experimental, its history is as short as the history of experimental psychology.

Two major figures of modern psychology and physiology have furnished original insights of crucial importance in the early development of stimulus-response positions. These two men were Pavlov (1927) and Thorndike (1933). Pavlov's use of experimental methods in identifying the important variables involved in a conditioning situation served as a major impetus to the developers of S-R theory. His work is well enough known not to require description here. Although none of the recognized contemporary theorists could fairly be described as a Pavlovian, these early experiments have persisted as major points of reference for theoretical clarification and induction. The work of Thorndike has also been of great importance, and the attempt to integrate his thought within the Pavlovian tradition has been a fruitful area of controversy and advance. The major current theoretical systems have in one way or another attempted thus to integrate the work of Pavlov and Thorndike. An excellent picture of the present thinking on this topic can be found in Postman's (1947) recent analysis and evaluation and in Meehl's (1947) comment.

The attempts at integration of Pavlovian and Thorndikian thinking have progressed as an area within the changing outlines of the behaviorist movement, which has itself expanded to include much of present American psychology. Spence (1948) has stated that behaviorism as a school no longer exists but that some of its early emphases are now almost ubiquitous. The first of these emphases is on what Watson (1913) called the "naïve use of consciousness." Watson stated that psychology did not need to deal with consciousness in the sense of observing special "states of consciousness." It should return, rather, to a nonreflective and naïve use of consciousness, which is the "instrument or tool with which all scientists work." A second important influence is the emphasis upon behavioral referents for concepts. A sample of this state of affairs is Bronfenbrenner's (1953) statement, in his current review of the field of personality, that operational definition is the eventual resting place of all good scientific concepts. Bronfenbrenner goes on to emphasize, however, the real point of present dissension: that the immediacy of the need for operational definition is variously interpreted.

Another source of energy for modern S-R theory, at least recently, has come from the field of philosophy of science and modern relational logic. The dual emphasis of Vienna logical positivism and Jamesian pragmatism on an empirical theory of meaning has helped furnish a broader base for thinking in S-R psychology. The influence of modern symbolic logic may be seen by glancing through recent volumes of the *Psychological Review* (e.g., Feigl, 1945; Koch, 1941; McReynolds, 1950; Moore, 1952). S-R theory has provided some outstanding examples of explicit and testable operational theory construction (Hull *et al.*, 1940). Whether or not the current "models" are long retained, the fact of their existence and use has pointed the way toward a general approach of importance to psychology.

---

*The author wishes to acknowledge the ideas and other help given him by Elisabeth C. Lambert.

Among recent influences in the development of S-R theory has been the insistence of Gestalt and Gestalt-type theorists on a larger unit of analysis. This has clearly contributed to a more intent appraisal of the definition of such root terms as "stimulus" and "response." Freudian psychology, too, has been of increasing interest to S-R theorists. Freudian speculation on the nature of motivation has led to the inclusion of much of the nonteleological area of McDougall's thinking, and interest has been growing also with regard to the Freudian mechanisms. The common aspects of the Darwinian tradition, accepted by both groups of theorists, has perhaps helped here. As Bergmann (1943) has pointed out, however, while the Freudians have accepted almost all of functionalism, S-R theory has been more cautious, because of its preoccupation with the behaviors of men and animals rather than only with their accomplishments. The Freudian emphasis on structure has also been approached warily. Structure analysis is essentially an emphasis on full description of a momentary state wherein a large number of dependent variables are described in interaction. S-R theorists have maintained their analysis of laws of development of partial structures as a prior commitment. More recently, and with increasing confidence in their experimental and theoretical tools, some S-R thinkers have emphasized structure analysis. The work of Miller (1951a) and of Whiting (1948) on conflict are some foremost examples of this trend. In general, the phenomena and insights of Freud, as well as the structural notions of Lewin, are being given more interested attention.

Zeal for knowledge which has practical consequences has been a continued motif in the tradition of behaviorism. Watson's programmatic boast about rearing children closer to the heart's desire of the parents is symbolic of an interest in *knowledge for use*, which has taken on new and more sophisticated forms in recent years but which has maintained its strength. And, of course, the eventual criterion for deciding theoretical issues may be seen to be the reinforcements provided to the society which underwrites the scientific endeavor. However, although any S-R scientist may be pleased by being able to make practical recommendations for training programs, like Jenkins and Stanley (1950), and although everyone enjoys reading Guthrie's quotations from a bird-training manual in a discussion of his theory, theirs is a different attitude from that of continued emphasis on particular practical problems. Whereas application of theory is desirable, the focus of the search is knowledge and understanding rather than immediate application. Selection of problems for study has been dictated chiefly by theory, rather than by immediate pressures of applied problems. This attitude is represented by F. H. Allport's (1942) recommendation to S.P.S.S.I. during World War II, that the need was not for values, which we had, but for a careful scientific analysis of the facts.

These statements apply to those S-R theorists who have contributed to social psychology, as well as to others. The work of the former has been aimed at the immense problem of finding the classes of data, and the conditions thereof, in general human (and animal) interaction which test the empirical laws and principles of S-R psychology, as previously generated experimentally. This work involves, among other things, utilization of data obtained by other social scientists, together with occasional or frequent laboratory clarification.

Recently there has been the very encouraging development of mutual attacks on the problems of social psychology by combined efforts of psychologists and other social scientists, such as anthropologists, economists, sociologists, and educators. In the forefront of this movement have been several S-R theorists (e.g., Miller and Dollard, 1941; Whiting and Child, 1953). Some theoretical intertranslation has been accomplished, and the notion of a general theory of social behavior remains an impelling idea, although the development of a multitude of "small systems" will no doubt continue to represent the more typical approach to this problem.

At present social psychology may be characterized as showing two opposite, although not dissonant, tendencies. One is the urgency of expansion to new problems and areas of data, and the other is a tendency toward systematization, consolidation, and refinement. MacLeod's (1947) emphasis on the phenomenological approach is, in part, an instance of the first tendency, an insistence on freedom of inquiry in the search of new problems. This mood is reflected by the interdisciplinary and intertheoretical advances mentioned above, and it also shows a belief in the desirability of a more comprehensive social theory and an expectation that personal experience still holds the secret of many new variables and theoretical notions.

The second tendency, toward refinement and systematization, also seems widespread. S-R theorists appear to have joined in each of these tendencies, wishing to maintain the formulation of S-R theory, and also showing increased interest in the phenomena and hypotheses of other researchers.

THE THEORETICAL SYSTEMS

Stimulus-response theory has developed along several lines, and the present discussion will furnish an introduction to the theories of Hull (1943, 1950), of Mowrer (1950), and of Guthrie (1935, 1952), which are outstanding examples of three major kinds of S-R theory. We shall call these systems "theoretical systems" to distinguish them from the "theoretical-empirical systems" which will be discussed in the next section. "Theoretical system" is a term which refers to the conceptual scheme as worked out by the theorist, including specifically the experimental reference data which serve to define the empirical realm which the conceptual scheme subsumes or explains. "Theoretical-empirical system" refers to a generalization of the conceptual scheme (that is, its concepts and the relations among the concepts) to some *additional realm* of data other than that which has helped define the theoretical system. We shall use the term here to refer to attempts to apply such theories as those of Hull, Mowrer, and Guthrie to areas of data which usually are considered as being in the field of personality and social psychology. It should be stressed that this is a different endeavor from that of testing implications of the theories within the area of the defining data itself. There is a broad history of studies of this latter kind, and we shall merely refer the reader to other discussions of this important problem (see Hovland, 1951; Brogden, 1951; Spence, 1951; Miller, 1951b; and Hilgard, 1951).

The distinction between these two kinds of system, the theoretical and the theoretical-empirical, is used here as a heuristic device. It is very difficult, in fact, to state at just what point a theoretical-empirical system has occurred. Further, the distinction is often not necessary, since some scientific questions, if correctly asked, can be answered by looking to the reference or defining data itself. One can ask, "Do 'values' affect perception?," and perhaps get an answer, "Yes," from one of Bruner and Postman's (1947) experiments. These experiments can also be used, in conjunction with other data and a conceptual scheme, to attempt an answer to further and more difficult questions (Bruner and Postman, 1949). It is this second use of data which is our present concern.

Perhaps we can best orient the discussion by taking an example of the behavior of an American child and directing our questions at this behavior. Our focus will not be, however, on the structure of the child's personality as an individual nor specifically on the cultural or social factors within which this behavior can also be viewed. Our focus is rather on the genesis of the behavior involved, and those aspects which we shall abstract for discussion will be pertinent to specific problems of learning theory. The experiments to which we shall refer deal with problems which are theoretically relevant and which appear requisite for scientific description of the process of learning.

A LEARNED BEHAVIOR SEQUENCE

Children are trained to do certain tasks — to keep themselves clean, to do chores, to read books, to write their names, and to like the "right" objects, people, and ideas. Any of these could furnish an example for us. Let us choose, however, the following situation in the life of an American child. We shall describe this situation, inserting reference letters, as follows:

> Johnny is three years old. His daily behavior during the summer months has come to include a certain sequence of acts which we shall describe here. Let us look at Johnny on a day when his learning of this sequence of acts is fairly well advanced. We note that Johnny is standing on a chair in the kitchen and has just reached into the cupboard for a box of cookies when the bell of the ice-cream truck rings (A). Johnny hesitates for a moment and then says, "Oh, darn!" He gets down quickly and runs immediately into the living room where his mother is sewing (B). He asks her, "Mommy, can I have a nickel to get some ice cream?" (C). His mother looks up from her work, smiles and says, "Go get a nickel from my pocketbook." (D). Johnny runs to the pocketbook and extracts a nickel (E), then runs out to the ice-cream truck, hands over his nickel (F), and with mouth watering (G), gets his ice cream (H), and eats it (I).

A CATEGORIZATION OF THE SEQUENCE

This description represents a common-sense description of Johnny's behavior. The letters have been placed in the description for the purpose of later analysis into categories according to S-R learning theory. We shall make our categorization of Johnny's behavior and then outline those aspects which different theorists have emphasized. The categories are the following:

(A), the sound of the ice-cream bell, is a *conditioned stimulus* (CS) which the remainder of the behavior sequence follows. (B), running into the sitting room, (C), asking for a nickel; (E), running and getting the nickel; and (F), running to the truck and handing over the nickel — all these can be termed *responses* which are *instrumental* and which are in a particular *sequence* to a *goal*. In (I), eating ice cream, a response is made which appears to terminate the sequence and provides *reinforcement*. It is to be noted that (D), mother's statement, is another *stimulus* to Johnny which we shall call a *subgoal*. An additional subgoal which should explicitly be mentioned is the nickel, which is a *stimulus* or stimulus object. Further, let us call (G), the watering of Johnny's mouth at the sight of the ice cream, a *conditioned response* (CR), in the sense that it once occurred only when ice cream was on Johnny's tongue or was being eaten, but now occurs sometime in the temporal interval between (A), the sounding of the truck's bell, and (H), the taste of the ice cream, which later we shall term the *unconditioned stimulus* (US). One other category is needed at this point — the *drive*, which motivates, at least in large part, this entire sequence of behavior. Since it is 4:30, we shall simplify our problem by assuming that it is hunger which is motivating Johnny, although we know that the "appetite," or "secondary drive," for ice cream is probably also present and active.

We have now presented a sequence of behavior in common-sense terms and have given names to some of the occurrences in this sequence. It must be kept in mind that we are not naming everything which goes on in this complex behavior, nor have we given names which would satisfy all the theorists whose ideas we wish to introduce.

There are differences among theorists in modern learning theory in regard to the following factors: (1) some would feel that not all of our categories are necessary or particularly fertile; (2) some would choose different terms for the categories on grounds of the "meaning" with which they wished to endow them, or on the grounds of their feeling that the rules of interrelations among the categories are different from what is here assumed. This last kind of difference among theorists may be due to (a) lack of clarity or explicitness of some categories and rules; (b) lack of evidence, or of ways of obtaining evidence at any particular time on specific issues; (c) differences in what would be accepted as evidence.

Before going on to our theorists, however, let us point out some things about what Johnny has learned. At some time in his life Johnny did not know that the sounds of the truck's bell, or nickels, could be connected with eating ice cream. At the time when we described his behavior, however, these stimuli (conditioned and subgoal stimuli) have been tied up with "unconditioned stimuli," or "reinforcing states of affairs," and have taken on new properties relevant to Johnny and his behavior. Further, we may note, if we look very carefully at Johnny (and *if* current theories are fitting), that his mouth will water when the stimulus of the bell is heard by him, i.e., the "unconditioned response" moves back in the chain so that Johnny can be said to behave as if he anticipated or "expected" ice cream when the bell sounded.

As this procedure goes on day by day, we may see that Johnny might tend to "short circuit" the sequence. Parts may drop out, or the instrumental acts near the goal may occur earlier in the sequence, as when Johnny goes to the pocketbook and takes the nickel without asking his mother. Further, if we measure the various instrumental acts which Johnny goes through, we may discover a "gradient of locomotion" in the sense that as Johnny approaches the goal, the speed with which he performs each act becomes relatively faster. Again, if we should use Bruner's "egometer" (Bruner and Goodman, 1947) to measure the apparent size of the nickel to Johnny before and after it has become tied up with Johnny's needs or his goals or both, we may find the experience has led to an accentuation for him of the size of the nickel, or that a perception has been learned or altered by the experience of adding this sequence of behavior to his repertoire (Lambert, Solomon, and Watson, 1949).

We say that Johnny "has learned something," since the ice-cream truck sounds have become important to him. What has he learned? How did he learn it? What are the laws necessary to predict the resultant or consequent behavior which we have described and named? What has changed in Johnny? These are some questions to which learning theorists attempt answers.

For our present purposes it will be necessary first to select from the richness of the theoretical systems of three kinds of learning theorists the most characteristic aspect of their positions and, second, to analyze in more detail what these theorists say about how behavior sequences like Johnny's are acquired and lost, and what behavior should characteristically accompany such acquisition and loss.

ADDITIONAL PROCESSES OF LEARNING

The experimental study of both Pavlovian conditioning and Thorndike's problem-solving situations has provided *several repeatable phenomena* which have been considered by most learning theorists and investigators as important kinds of data for any learning theory to deal with. Let us discuss these kinds of data in the context of our example in order that we may better see the problem that a systematic theorist faces in integrating this area. The uninitiated reader is referred, for a more detailed discussion than follows, to the several definitive chapters in the *Handbook of Experimental Psychology* (Stevens, 1951) on the broad empirical base of learning experimentation, particularly concerning the phenomena which have been of special interest to S-R theorists.

Although until recently most learning theory controversies (see Spence, 1951) have been focused around how to state the necessary and/or sufficient conditions for learning, there are some signs that other (subsidiary) empirical and theoretical principles are of equally great importance for work in social psychology. The notion of "stimulus generalization," for example, may provide hunches for défining "modal personalities," and for assisting in the tremendous problem of categorizing social situations (see Tolman, 1951, and Parsons and Shils, 1951). We shall conclude our discussion of the theoretical systems by summarizing the current systematic positions of these theories, and by adding certain further concepts which have not yet been fully integrated into the systems. This will provide us with a clear picture of just what is at our disposal for use in more purely social psychological areas.

Temporal relations. Conditioned responses are most strongly established when the unconditioned (or reinforcing) stimulus follows the conditioned stimulus by a short time interval. With human subjects, the strongest conditioning was found by Kimble (1947) to occur when the conditioned stimulus preceded the unconditioned by 0.5 second. The empirical facts seem to confirm the general notion that in the classical conditioning situation there is a relationship of decreasing strength of conditioning as this temporal interval is increased. Backward conditioning, where the conditioned stimulus follows the unconditioned stimulus, has been reported (see Hovland, 1951) but is generally less effective than when the conditioned stimulus precedes the unconditioned stimulus. These temporal relations, as well as others, have been important in theorizing about learning since the time of Pavlov.

Spontaneous recovery. It appears to be the case with most conditioned responses that when time is allowed to elapse after the response is extinguished, the response may reappear spontaneously and increase in strength without further reinforcement. In other words, if Johnny's mother desired to end the behavior we have described above and had apparently succeeded in doing so, the chances are that the responses would tend to recur at some later time. Such spontaneous recovery provides an interpretation for much of the seeming naughtiness of children, who, having learned the right thing to do after much trouble, occasionally revert to a bad habit.

Stimulus generalization. One of the focal points, and a remaining problem, in conditioning and learning theory is the fact of stimulus generalization. This concept refers to the fact that when a response has come to occur to a certain stimulus or stimulus pattern, the same response will also occur to similar stimuli without any reinforcement for such occurrences. Hovland has recently summarized the salient fact:

> Immediately following the conditioning trials generalization may be almost complete, that is, as large a response being made to stimuli not previously paired with the unconditioned stimulus, as to the stimuli given paired presentation. With subsequent testing, the generalized responses appear to extinguish more rapidly than those established to the original conditioned stimulus. (Hovland, 1951, p. 616)

One of the important areas for quantification in learning theory is the identification of factors that affect the shape of "gradients" of generalization, but for general purposes one can imagine that the form is roughly that of a falling off of the amplitude of the response as one makes the conditioned stimulus less and less similar to the conditioned stimulus that had been paired with the unconditioned stimulus. Thus if we varied the sound of the ice-cream bell up and down the intensity scale away from the tone to which Johnny has been conditioned, we should find Johnny's salivating response decreasing in amplitude. The shape of any generalization gradient is, however, a difficult matter to predict at the present time except under limited conditions where psycho-physical controls can be maintained. Lawrence (1952) has recently raised some serious questions on this score. As we shall see below, gradients of stimulus generalization have been studied in many kinds of situations, and the general concept has been found useful by such varied social scientists as Stroud (1946) in educational psychology, Miller and Dollard (1941) in social psychology, Dollard and Miller (1950) in personality theory, and Murdock (1949) in cultural anthropology. The knottiest problem in this area involves first how to measure *learned* stimulus generalization gradients and, second, a statement of the complex laws of how such continua or scales become learned by the individual.

Discrimination. The important fact of discrimination is stated by Hovland as follows:

> If reinforcement is always present when the original conditioned stimulus occurs but is always absent when the generalized stimulus occurs, the combination of increased strength of the conditioned response and extinction of the generalized response reduces the degree of generalization until finally the response occurs to only a very narrow range of stimuli. The accuracy of discrimination is reduced when the stimuli are presented rapidly and in irregular order. (Hovland, 1951, p. 615)

Therefore, although in the early stages of learning the sequence of behavior Johnny might have responded to many kinds of bells and cymbals, he would soon come to respond by asking for a nickel only when the bell was the bell of the ice-cream truck, or one very similar. This would be the case unless, of course, his mother always gave him a nickel when any bell rang (an improbable situation), and he had previously learned to use nickels to go to the store.

Higher order conditioning. Probably the notion of higher order conditioning is the most important concept which conditioning experimenters have provided for understanding complex human social behavior. The term refers to the experimental fact that a conditioned stimulus which has previously been associated with an unconditioned stimulus, will, independently of any further reinforcements, in some circumstances function as an "unconditioned stimulus" in setting up a second conditioned response to a third stimulus. Such conditioning can sometimes be continued for several stages. This capacity for previously neutral stimuli to take on the capacity of being reinforcers is a core notion in learning theorists' approach to the development of valuing behavior towards environmental objects. In our example, ice-cream bells, five-cent pieces, and probably the sight of ice-cream vendors have all come to "have value" to Johnny. Furthermore, Johnny would probably learn new sequences just to get nickels and go out of his way just to get close to sounds of ice-cream trucks.

Anticipatory phenomena. The flow of saliva which follows the sounding of the conditioned stimulus in Johnny's behavior sequence originally occurred only when the ice cream was in his mouth. This "fractional anticipatory goal response" has, in learning, commenced to occur out of place in the sequence, so that Johnny may be said "to anticipate" getting ice cream. Such anticipatory responses can be seen in some kinds of experimental situations; in others, they cannot be seen directly but have been inferred or "hypothecated" by some experimenters (see Spence, 1951).

If nursery school children are taught to enter a room and turn a crank and get candy, the saliva flow which occurs while they are working for candy may require the use of a mop for the floor, and such children are quite free in calling loudly, "I want some candy" or "I came here to get candy," providing clear evidence of their "anticipation" (Lambert, 1950).

Such observations of child behavior are very silly unless the major point is grasped; such "purposive" behavior is so much with us that some "cognitive" theorists have tended to *assume* it in their theorizing about human beings or about rats. S-R theories have viewed the problem of purposes as *one which is not to be assumed,* but as one which *must be explained* in the sense of outlining the laws whereby organisms develop those "purposes" they are not born with. Further experimentation with anticipatory phenomena may help in this matter.

These, then, are some of the basic facts of conditioning. We have here limited our discussion to repeatable phenomena which have been obtained with human subjects. Many of the definitive experiments on these phenomena have, however, been conducted with infrahuman organisms as subjects. This has been the case for at least two reasons: (1) some psychologists are interested mainly in comparative psychology and the study of infrahuman behavior; (2) and most important, infrahuman experimental situations permit much more careful control of the broad scope of variables which must be taken into consideration in these kinds of studies. The aim of research in the second case is often the attempt to uncover more clearly some of the basic factors in the environmental control of behavior changes irrespective of species. In the present context this is our major interest, although the study of comparative social psychology (Smith and Ross, 1952) is in itself an area which probably will increase in importance as social psychologists begin to take into consideration more seriously the need to control genetic factors in their experiments (see Tinbergen, 1951).

It is also important to note that a great many of the above relationships have been found in other kinds of learning situations than those of classical conditioning, and with human beings as well as lower organisms. Skinner (1938), for example, has been able to obtain highly repeatable and clear-cut data in all of these areas by using an instrumental conditioning situation similar to Thorndike's problem box.

THREE THEORIES OF ACQUISITION AND EXTINCTION

Let us quickly outline the major characteristics of the three types of theory concerned with *what* is learned and *how*. We shall also try to show how these theories deal with the categories as named above in our analysis of Johnny's sequence of behavior.

"S-R contiguity" theory. The most outspoken exponent of this type of theory is Guthrie (1952). Guthrie talks in the language of responses and of the contiguity between them and stimuli (internal and external). However, he would suggest that we watch the development of Johnny's sequence of acts much more precisely than we have as yet done. In fact, hardly anyone has ever watched an organism sufficiently carefully to give referents to Guthrie's more "molecular" terms.

Guthrie would suggest that the "acts" of Johnny developed out of tiny, more or less discrete *movements,* which become attached to *stimuli* (say the internal physiological stimuli of Johnny's drive state). These stimuli evoke and maintain the sequence of movements by laws of strict contiguity. The notion of contiguity has been restricted by Guthrie to the meaning, "temporal simultaneity." The simple and sovereign principle of learning here, then, is that if a stimulus has occurred in contiguity with a response, the response will always so occur, unless some other response has become conditioned to the same stimulus. Getting the ice cream and eating it is merely a "terminating response" which affects some stimuli (say, the drive state) so that new and interfering movements cannot become conditioned to the stimuli prior in time to such ice cream eating.

"Extinction," or disappearance of the response, is a matter of *interference,* where the necessary abrupt change in stimuli or responses does not occur (for example, Johnny keeps on desiring ice cream after having been out at the truck) and new and interfering responses become attached to the stimuli. Johnny, for example, might leave the truck and sit down on the porch and mope. "Moping" responses might become conditioned to the internal stimuli representing the desire for ice cream.

Guthrie really suggests four specific ways in which extinction can be accomplished: (1) the conditioned stimulus may be acting, but since it is below threshold the response will not occur; (2) the response may be eliminated through exhaustion or fatigue; (3) the response may be inhibited by the action of incompatible responses. These three ways "involve the presence of the cue and the prevention of the response" (1952, p. 59), and the reader can readily apply them to the situation with Johnny. (4) The fourth method Guthrie describes is a case where "the substitute — conditioned — stimulus is applied so soon after its response that the response is still in what is called its refractory period." In principle, then, if we rang the ice-cream bell directly after Johnny salivated (assuming "salivating" is the response conditioned to the bell), then some new response would become conditioned to the bell, and if salivation has been occurring early enough in the sequence, some new sequence of behavior will develop. Of course, salivating may be only apparently conditioned to the bell, when in actuality some other response (to which salivating is attached in a chain) may be the key to the sequence. In this case the sequence might continue as before unless we diagnosed correctly which response is the "key" one to the sequence which costs the mother a nickel, and then applied this rule to the refractory period of this key response.

"S-R reinforcement" theory. The most elaborate and formally systematic theory of learning is that of Hull (1943), which emphasizes the occurrence of drive reduction as a prime mover of learning. Hull's formulations represent an attempt to integrate the Pavlovian classical conditioning tradition with the reinforcement or law of effect tradition stemming from certain British associationistic philosophers, and more recently from Thorndike (1933). Thorndike was interested in the temporal relationship of the instrumental response (like Johnny's act of the receipt of the reward (or punishment). The handing the nickel to the ice-cream man) and strength of the learning bond between a stimulus and a response was found to be some function, other things being experimentally equal, of the temporal (or spatial) discrepancy between the response and the reduction of some aspect of the drive state.

Hull also considered theoretically the importance of the temporal contiguity of the conditioned and unconditioned stimuli, and anchored to the reinforcement context the effects usually ascribed to such asymmetry of stimuli. He did this by broadening Pavlov's notion of the physiological trace left by the CS, in order to tie such a trace up with the occurrence of the "reinforcing state of affairs." Such traces decay with time and are less strongly conditioned to the instrumental response the more it is separated in time from reinforcement.

In short, Pavlovian conditioning phenomena, as well as simple problem-solving phenomena, become tied to the same basic event — the reduction of a drive or strong impelling stimulus — which Hull calls a reinforcement. Extinction is explained by Hull in terms of a dual process. If Johnny indulges in his sequence and it is not reinforced, inhibition will develop. One kind of inhibition is close to the common-sense notion of fatigue which develops just because Johnny does the act, and the effects of which dissipate fairly quickly with time; the second kind of inhibition is permanent but develops more slowly. Any "resting response" which reduces the fatigue-type inhibition (another form of reinforcement) leads to the increase in probability of the resting response becoming attached to the cues of the ice-cream bell, so that Johnny dawdles or sulks, instead of running to mother for the nickel for the ice cream.

A separate notion that Hull has taken from Pavlov, and has altered, has to do with "secondary reinforcement," where a previously neutral stimulus (like the nickel in our example) comes to take on the capacity to reinforce responses to other stimuli. The nickel thereby becomes a reward to some degree and comes to function as such — Johnny would learn new sequences of responses in order to get nickels (after nickels have taken on secondary reinforcing properties).

After analysis of the general experimental literature in learning, Hull attempted to set up a "law of reinforcement" which would state the necessary and sufficient conditions for learning to occur. To do this he found it necessary, of course, to specify, in the light of available evidence, his best guesses as to what the other major factors are that one must relate to the basic law in order to increase its utility to experimentation and to generalizing. Since most of his guesses were based on certain defining evidence present when, in 1943, he first proposed this major theoretical system, he found it felicitous to change the system in some respects in 1950. In light of the fact that a good percentage of basic research in learning was focused on testing in various ways his early formulation, he managed, until his premature death in 1952, to achieve a great degree of cooperation from other people in his theoretical endeavor, and to receive free (and often rather strong) suggestions from others. We shall discuss Hull's system more thoroughly after a discussion of our third type of theorist.

Dual theories of acquisition. It may be of value to discuss in the context of Johnny's behavior another notion of what is involved in the process of acquisition of behavior sequences. Several modern S-R theorists have come to the conclusion that two principles are necessary to handle the experimental literature in the fields of conditioning and problem solving. These dual theorists include B. F. Skinner (1938), Harold Schlosberg (1937), O. H. Mowrer (1950), and Eleanor J. Gibson (1952). These theorists accept both the principle of contiguity stimuli and some principle of reinforcement (that is, success, but not always drive reduction) as necessary in formulating an adequate theory of learning.

Skinner's (1938) famous distinction is between "type R" conditioned responses and "type S" conditioned responses. In type R conditioning, which depends upon the principle of reinforcement, the early occurrences of the conditioned (instrumental) response are not under the control of the experimenter. It is as if the experimenter were the ice-cream man, who cannot force Johnny to bring him a nickel for ice cream but must wait until a stimulus brings the first act out. Of course, once the act is out, the ice-cream man has a modicum of control by virtue of the degree to which he has reinforced Johnny's behavior. In type S (classical) conditioning, the experimenter can control the occurrence of the to-be-conditioned response, just as Pavlov could bring about the dog's salivation by introduction of food powder in the dog's mouth. According to social convention, Johnny's mother, if not the ice-cream man, might be permitted to bring about salivating responses in Johnny by presenting him with his first taste of ice cream along with his other food at meal time, whether Johnny wished it or not.

Skinner suggests that conditioning of type R depends on the occurrence of a reinforcing stimulus (not necessarily drive reducing) following the response. Type S conditioning depends on the temporal relationships the experimenter (or mother) sets up between the CS and the US. In most complex behavior sequences, like Johnny's, both kinds of conditioning occur. The probability of the salivary response (type S conditioning) might be determined by the temporal contiguity of the bell and the receipt of ice cream. The other (type R) response chain (running and asking mother for a nickel, etc.) might depend essentially on receipt of ice cream after the responses have occurred. Skinner emphasizes the fact that in type R the experimenter often has no notion as to what stimuli evoked the first response. That is, the experimenter is in the position of the ice-cream man who sees Johnny come running out for the first time and doesn't know who was the source of the nickel, or whether Johnny had received his mother's permission or not. This is, in a sense, the problem of the Rorschach tester who gets a great many responses from the subject but has little notion as to what evoked them except that the ink blot and the results of the individual's past experiences have something to do with it. Skinner has also suggested that type R responses are largely skeletal, whereas type S responses involve those types of effectors that are controlled by the autonomic nervous system.

Mowrer's (1947) conception is close to that of Skinner, as is that of Schlosberg. Mowrer has attempted to show that anxiety can be profitably viewed as a secondary drive and anxiety reduction as a reinforcement. Learning to be "anxious" occurs according to laws of stimulus contiguity, whereas there must be reinforcement (drive reduction) for a skeletal type of response to be learned to the anxiety stimulus.

There is no data yet which can crucially decide the issue as to whether an explicit monistic theory of either the Hull or the Guthrie type is more or less adequate than these dual theories. White and Schlosberg (1952) have reported evidence which they feel is opposed to either the Mowrerian or the Hullian interpretation, but more evidence of this kind will be necessary for these issues to be clearly decided.

We now turn to a more systematic presentation of the theories of Hull, Mowrer, and Guthrie.

HULL'S THEORETICAL SYSTEM

Leeper, who is critical of Hull's theory, has summarized Hull's importance as due to the fact that "more than any other psychologist, he has emphasized that the goal of psychology is the discovery of quantitative relations among basic variables" (Leeper, 1952, p. 490). Leeper also agrees with Hilgard's (1948) statement to the effect that work in the field of learning will need ultimately to follow the *form* which has been suggested by Hull's work — we must look to the development of abstract, quantitative propositions stated to a considerable degree in terms of inferred factors (Leeper, 1952). Although, as we shall see, new methods of theorizing are already developing in learning theory (Mosteller and Bush, 1951), undoubtedly Hull's preliminary formulations have greatly speeded up theoretical development in the field — some would say (Skinner, 1950) that they have speeded it up *too* much.

There are three recurrent themes in Hull's thinking: (1) Vulnerability of one's theory is vital to scientific progress (that is, the propositions of the theory *can* be tested and proved wrong), and explicit mathematical formulation, even when inductive procedures are not elegant, is one means of assuring vulnerability. (2) The hypothetico-deductive method, where the hypothetical relationships are based on the "best" induction possible, and where rigorous deductions from the set of variables can be tested, is a defensible procedure and may help to advance the integration of the behavior sciences. (3) Any theory should be expected to change as crucial experiments are carried out, and as the theory is generalized to new areas of data. Hull hoped, however, that his theory, which attempts to integrate the experimental findings in the areas of conditioning and problem solving, would provide at least a first approximation to the variables needed in the analysis of more complex phenomena.

An example of Hullian analysis. We shall use Hull's (1943) original model, in the main, in this exposition, since it is this statement of Hull's position which has given rise to a whole movement in experiment and theorizing. Hull (1950) has recently outlined certain changes in this 1943 theory, but the present exposition will provide an adequate introduction to this outline. Hull's death has left future development of his theory in the hands of others, and further alterations are to be expected from his original statement of theory. In this section we shall present the Hullian position as it emerges from a recent attempt by Brown and Farber (1951) to formally add "emotion" as a new intervening variable in the system. Their presentation has two main virtues — it is compact, and it shows clearly the susceptibility of the theory to change, where desirable.

First, let us look at Brown and Farber's reasons for adding another construct. After considering several sources of evidence and thought about "emotion," and considering Mowrer's theorizing of a decade ago (see Mowrer, 1950) and Leeper's (1948) more recent suggestion that emotion be analyzed as a motive rather than as a disruptive state, they attempt to integrate these notions around the consideration of "emotion" as an intervening variable. The notions involved in their induction procedure stem from the work of Mowrer (1950) and Miller (1951b) on fear, the thinking involved in the Yale group's frustration-aggression hypotheses (Dollard *et al.,* 1930), Maier's (1949) analysis of frustration and "abnormal fixations" in rats, and from the work of French (1944) on frustration from the Lewinian point of view. These various sources led them to some ideas on how the emotion of "frustration" (or "anger") can be placed within Hull's system, assuming as they do that:

> A theory of emotion, stripped of its nonessentials, is simply a series of guesses as to the manner in which emotions are presumed to result from environmental events (past and present, internal and external), guesses as to how they may resemble, differ from, and interact with other hypothetical states or processes within the organism, and guesses as to how different emotions, or different degrees of the same emotion may affect behavior differentially. It is furthermore stipulated that these various guessed-at relationships be stated specifically enough to assure their susceptibility to experimental test. (Brown and Farber, 1951, p. 468)

Let us use the "whole-teaching" method of explication and ask the reader to look at Fig. 1, which will provide us with the results of Brown and Farber's thinking as well as much of Hull's theorizing. Let us first note that the analysis moves from left to right across the page. Humorously speaking, this may be the reason why Hebb (1951) and Bronfenbrenner (1953) refer to rigorous theorists as the "rightists" in contemporary psychology.

In the left column are the statements of the "stimulus variables" of the system, i.e., the antecedent manipulable conditions, which, in conjunction with the intervening variables (middle column), determine the "observable response characteristics." Such antecedent-to-response analyses are generally characteristic of S-R theory. Skinner (1945) has suggested that this be considered the hallmark of "behavioristic" analysis. We shall discuss later some of the other characteristics of S-R theory.

Before focusing on the star of this theoretical performance (i.e., the "F" in the middle of the diagram) let us look at Hull's intervening variables that are present in the analysis. First,

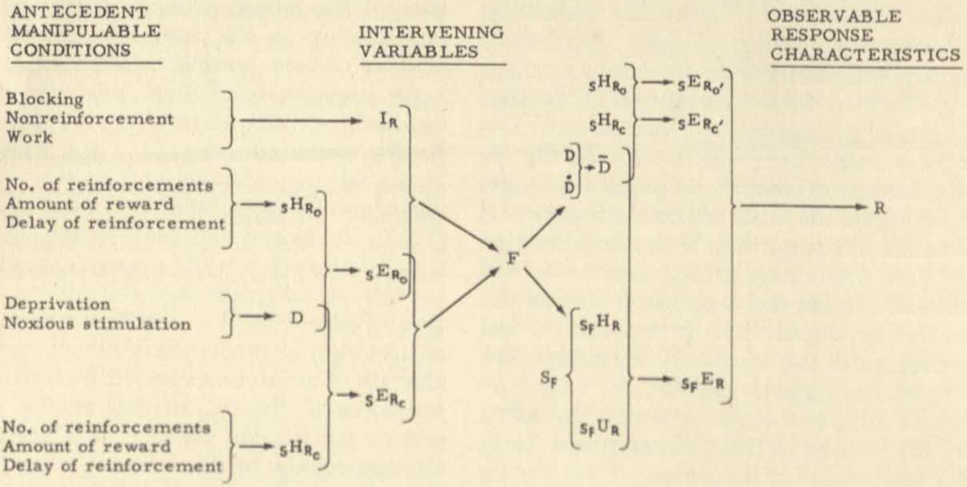


FIG. 1. Diagram summarizing the antecedent manipulable conditions and the hypothetical variables employed in the formulation of the present theory of frustration. The symbols for the hypothetical variables are defined as follows: $_SH_{R_o}$ = habit strength of an ongoing response tendency; D = relevant drive; $_SH_{R_c}$ = habit strength of competing (weaker) response tendency; I_R = competitive inhibitory tendency; $_SE_{R_o}$ = initial excitatory strength of ongoing response tendency; $_SE_{R_c}$ = initial excitatory strength of competing tendency; F = frustration state; $\dot{D}$ = irrelevant drive due to frustration; $\bar{D}$ = effective drive strength resulting from combination of D and $\dot{D}$; S_F = frustration-generated stimuli; $_{S_F}H_R$ = acquired habits associated with frustration; $_{S_F}U_R$ = innate response tendencies associated with frustration; $_SE_{R_o'}$ = final excitatory strength of initially dominant response tendency; $_SE_{R_c'}$ = final excitatory strength of initially competing response tendency; $_{S_F}E_R$ = strength of excitatory tendency (either innate or learned) associated with frustration. (From Brown and Farber, 1951)

note that this diagram is, like a Lewinian diagram, a "picture" in symbols of a very short molar sequence of behavior. That is, it is an attempt to capture a quantifiable picture of an organism in an act of being "frustrated," which involves being "in conflict." If one considers the diagram as including only the intervening variables and observable behavior characteristics, then the diagram is analogous to Lewin's "momentary field." Further speaking in Lewin's (1935) terms, these two columns set up the fuller specification of the relatively empty formula $B = (p, e)$.

One of the important *differences* between this kind of diagram and that of a field theorist is that the S-R theorist attempts to go further in setting up the antecedent manipulable conditions which elicit the particular quantified pattern of intervening variables and of observable response characteristics which define "the present field." This statement of "stimulus-intervening variable" laws is contingent, of course, upon the limitations of learning itself. That is, certain factors which are responsible for behavior change (e.g., fatigue processes and innate processes) are not fully "manipulable," and therefore must be "built into" the set of intervening variables. Theoretical models differ in the degree to which they rest upon such built-in factors as "prägnanz," "libido," etc., and in doing so assume that behavior is to differing degrees "docile" to environmental events.

We have assumed that Johnny learned the particular response sequence that we analyzed above. We do not mean that the process of learning substantively created his walking, talking, and salivating. We mean merely that as a result of the operation of certain antecedent conditions, Johnny came to behave in the way we described, and that as a result, in part, of certain currently operating conditions Johnny will either continue to go after ice cream every afternoon that the bell rings, or he will not. That is, he will "learn" (i.e., be "reorganized") as a result of what is happening now, so that the particular pattern of intervening variables will be a little bit altered as a result of his "experience" on this particular afternoon. Thus, dependent on the currently active environmental conditions of tomorrow, his behavior will be *different* tomorrow. "Learning" is happening to Johnny whenever the environmental conditions active on him while he is performing meet the requirements for learning to occur. The problem of learning theory is simply to specify the conditions under which changes in structures occur.

The major structural or "associationistic" units in Fig. 1 are the intervening variables, ${}_{S}H_{R_0}$ and ${}_{S}H_{R_c}$. The factors responsible for the particular strength of these at the present moment are given in the list of antecedent conditions: number of previous reinforcements, amount of reward present in those previous reinforcements, and the delay in the reinforcement (i.e., the delay between the time of arrival in the previous situation (S) and the reward, or the delay between the occurrence of the response sequence and the reward in past commerce with this particular situation).

Let us return to our example at this point and view Fig. 1 as representing Johnny's state when the ice-cream bell rang. The reader should recall that Johnny was busily engaged in reaching for the cookie box at the time that the sounds of the ice-cream truck intruded themselves into the situation. This tendency to go after cookie boxes has a certain "habit strength" (${}_{S}H_{R}$) even when the ice-cream bell intrudes on the situation. Let us further suppose that this is the first time that Johnny has ever gotten around to dealing in this way with the cookie box. The habit strength is a result of Johnny's having symbolically recalled the time when he had pleased himself and amused his parents by climbing up on a kitchen chair to lick the remainder from a bowl in which his mother had been mixing cake frosting. The strength of this "associative" factor in Johnny's present behavior is referred to by ${}_{S}H_{R_c}$, which Brown and Farber describe as "habit strength of competing (weaker) response tendency." The habit strength of the major response tendency (getting ice cream) is referred to by the rubric ${}_{S}H_{R_0}$.

This general intervening variable, ${}_{S}H_{R}$, is one of the major concepts in Hull's theory and, it is important to note, the concept is defined not only on the response side (that is, objective response characteristics to be expected with changes in ${}_{S}H_{R}$) but also on the antecedent side (that is, in terms of the antecedent conditions, a knowledge of which permits us to predict the strength of ${}_{S}H_{R}$). Our example should make it clear to the reader that the concept ${}_{S}H_{R}$ is a very general one which could subsume as special cases a large number of specific classes or instances of habits, such as getting cookie boxes or going after ice cream.

Let us now look at the symbol D in the first column of intervening variables. Brown and Farber describe this as the strength of the relevant drive. It should be noted that the brackets in the diagram suggest that there is a certain strength of this relevant drive for each of the response tendencies. Such a concept forces us to be explicit with regard to the concept of motivation. In our present example, we can tentatively name the drive involved both with ${}_{S}H_{R_c}$ and with ${}_{S}H_{R_0}$ as hunger. The strengths of the drive components involved in both responses are a function of the antecedent conditions of degree of deprivation — in one case "ice-cream deprivation," and in the other "cookie deprivation."

At the top of the diagram the reader will note that there is an intervening variable entitled I_R which involves what is, in effect, an additional competing factor which decreases the strength of what we have called the response tendency "to get ice cream." The strength of this "inhibitory" factor is a function of the antecedent conditions of the degree of blocking which the response tendency has very recently encountered, the degree to which Johnny has tried "to get ice cream" in this situation but failed (the degree to which nonreinforcement is operating) and, finally, the degree to which Johnny has done work in the present situation in his search for ice cream.

Having stated these major concepts in Hull's system, let us now look at how they are interrelated. Hull (1943) has suggested that ${}_{S}H_{R}$ and D are related in a multiplicative way to produce ${}_{S}E_{R}$. For example, in Fig. 1, ${}_{S}E_{R_0}$ is a resultant of ${}_{S}H_{R_0} \times D$. Furthermore, this resultant ${}_{S}E_{R_0}$ can be viewed as being in competition with either the I_R operative or the ${}_{S}E_{R_c}$ (or both) operative at a given time. Such competition between "inhibition" and "initial excitatory strength of a tendency" or between two "initial excitatory strengths of two competing tendencies" results in the arousal of the intervening variable frustration (F in the diagram). The two arrows which lead away from F refer to the two major effects which F has in the on-going situation. The upper one indicates that frustration (F) produces a certain degree of irrelevant drive ($\dot{D}$) which combines with the relevant drive (D) to produce the effective drive ($\bar{D}$). The second major effect of F is shown by the lower arrow. This refers to the fact that frustration tends to generate certain *specific* frustration stimuli (S_F). These stimuli, Brown and Farber assume, may have innately associated response tendencies (${}_{S_F}U_R$), or such specific frustration stimuli may have acquired new associative connections through past processes of learning (${}_{S_F}H_R$). The last column of intervening variables suggests how certain "final excitatory strengths" of the various response tendencies may be summated (algebraically) in order to produce one or more observable responses, the nature of which will depend upon the relative strengths of these several possible excitatory tendencies.

In our protocol of Johnny's behavior we reported that in his conflictful state, "Johnny hesitates for a moment and then says 'Oh, darn!' He gets down quickly and runs immediately into the living room . . ." The remark "Oh, darn!" (the *first* response tendency in Brown and Farber's theoretical formulation) can be interpreted as a very strongly learned response to certain specific frustration stimuli, learned by way of some identification mechanism (see below). Once this strongly learned and specific response had been emitted, then the stronger of the remaining response tendencies — getting ice cream — took over, and Johnny began the sequence of behavior we have recorded above. It is important to note that Brown and Farber's theory includes the important prediction that whichever responses occur in the "conflict situation" will occur with greater intensity because the frustration has increased the potential driving factor in the situation. Johnny "gets down quickly and runs immediately . . ." Hull has previously stated that the effect of an increase in drive on later behavior can be found in any of the following ways: (1) an increase in amplitude of the response, particularly if the response is an "emotional" one; (2) a decrease in the latency of the response; (3) an increase in the probability of the occurrence of the response; or (4) temporary increase in resistance to extinction. Such changes are also used to infer the effect of an increase in ${}_{S}H_{R}$ (other things constant), or an effect through any other factor which increases the quantity of the intervening variable ${}_{S}E_{R}$, since ${}_{S}E_{R}$ is the factor which is closest to the observable behavior.

Let us review at this point what we have done in explicating Hull's and Brown and Farber's theorizing. First, we have presented a categorization of certain observable aspects of Johnny's behavior in common-sense terms. We then translated these observable aspects into the categories utilized for descriptive purposes by various learning experimenters. Then we analyzed the theories of the processes of acquisition and extinction that have been developed by the S-R reinforcement theorists and by Guthrie's S-R contiguity theory. We then outlined the empirical facts involved in certain *other* areas of the process of behavior change, and finally we have presented an analysis of Johnny's behavior in terms of Hullian theory.

In doing this latter task we have presented only the major concepts which Hull's system uses, and these have been largely in the form in which he defined them in his important 1943 work. There are other intervening variables which Hull uses and which define his *total* system. Further, we have only cursorily dealt with the relationships Hull has postulated as obtaining among his intervening variables, and we have neglected the detailed mathematical statement of the relationships which hold, according to theory, between the "antecedent-condition variables" and the "intervening variables," and between the intervening variables and the consequent behavior variables. The reader is referred to other sources for most of these more precise statements (see Hull, 1943, 1951).

The nature of intervening variables. It is important at this point to raise the following question: What *are* these intervening variables? Do they exist anywhere, or are they just vagaries which Hull has foisted upon us? Two answers can be made to this question. Spence (1951), for example, and others have interpreted Hull's theory as a molar one, and they consider these intervening variables as "existing" only in the sense that the unseen atom exists, i.e., they are defined (operationally) in terms of antecedent "inputs" (observable, and for experimenters, manipulable) and also in terms of observable, measurable consequent "outputs." In fact, Hull's books are filled with references to such defining data. It is because of this fact that we have presumed to call Hull's theory a "theoretical system," whereas for contrast consider the "General Theory of Action" of Parsons and Shils, which has been termed a "categorical" system, wherein some educated guesses as to the important categories necessary for the analysis of action, and only such guesses as to categories, have been worked out. Although some readers might choose to place the intervening variables in some "phenomenal space," and in so doing might generate hunches for further study, this is not the intent of Hull. His intent is to leave the intervening variables in a mathematical space (as a factor analyst might do), and obtain deductions for test by way of mathematical manipulation in a "mathematical phenomenal space." A good percentage of psychological experimentation in the past few years has been devoted to such an attempt (see Spence, 1952b).

The second answer to our question about the nature of intervening variables has been suggested at times by Hull (1943). He hopes that these intervening variables may eventually be related to physiological processes which will be independently defined by physiological measurements. In fact, Hull has added "meaning" to such concepts as ${}_{S}H_{R}$ by trying to make a place for such definitions. In light of the papers by Klein and Krech (1951) and Hebb (1951), such matters may become increasingly important to social and personality psychologists, but for the present discussion we can consider them irrelevant. In this second answer to the question, we have dealt with some of the reasons why some people have called Hull's theory such names as "atomistic" because they prefer a different kind of physiology from that which he appears to espouse, or "reductionistic" because they are opposed to or uninterested in relating "psychological" intervening variables to the soma. According to Spence's interpretation, given above, such criticisms may be considered irrelevant.

Let us return for a final look at Hull's general theory and note an important distinction which Sears (1951) has emphasized. There are two foci which can be used when analyzing behavior in this S-R manner. In short, one can be interested in the changes of structure which occur (i.e., learning), or one can be interested in the on-going action itself (i.e., performance) and its consequences for others. We might be interested in what the experiences do to Johnny (learning), or we may be interested in the fact that he did *not* take the cookie box, or that children will do all kinds of things to get ice-cream cones. The only important moral here is, as we have stated elsewhere, that both these kinds of analysis are important, and S-R theory is increasingly grappling with both. F. H. Allport's (1940) "event system" theory can perhaps be viewed as an extension of S-R analysis into those aspects of performance which have hitherto been left largely in the hands of other social scientists.

The openness of Hull's system. One final point regarding S-R theory deserves statement here. Note in Fig. 1 the degree to which S-R theory is open to integration with other social sciences. "Social situations" can be viewed as the language area in which the learning and performance aspects of Johnny's action are developed and controlled, and therefore integration with concepts of "market" and "social setting" is very possible. The problems involved in such integration are great, however, since our sociological friends who, as Doob (1952) suggests, study the "stimuli" involved in everyday behavior, have their own ways of defining "social situations."

Integration of sociological and psychological thinking may come by way either of their definition of "sociological stimulus factors" in terms of some psychological theory such as Hull's, or through redefinition of psychological theory in terms of some sociological theory, such as that of Homans (1950). The probable long-range alternative is the developing of a kind of metalanguage which will permit a rough translation from one kind of theory to another. It may be, however, that other constructs in social science will take on the meaning of the terms of the most powerful theory extant at any given time. If so, the fact that one can defend Hull's theory as being an approximation to a theoretical system, whereas that of Parsons is confessed to be only a categorical system, may be an important consideration for social psychologists.

Hull's postulate system. The above discussion is only a rough summary of Hull's system. We have left implicit much of the content of Hull's postulates and corollaries. Yet it is to *this exact statement* of the system that the critic and/or generalizer must turn. Hull's (1950) new postulate set represents a consolidation and rearrangement of the earlier (1943) system. The most important change is a de-emphasis of the notion ${}_{S}H_{R}$ (or habit strength) and greater emphasis on ${}_{S}E_{R}$ (or excitatory potential), which is now the most important notion in the system.

Let us mention just two matters with regard to the postulate system. The first is that Hull's postulates include explicit recognition of almost all of the "additional processes of learning" which we have outlined above. Hull accepts all of these as, so to speak, important basic notions necessary to a predictive theory of learning. Guthrie, on the other hand, tends to consider these as phenomena which are to be explained by what we have termed his "principle of acquisition." The reader desirous of getting a rough notion of Hull's postulates can find it in our discussion of the additional empirical functions discussed above. The notion of Rg, or "fractional anticipatory goal response," is not systematically dealt with in either of Hull's statements of his system, but Spence (1951) has displayed the way in which it would be viewed as functioning in the system.

The second matter regarding the postulates which is germane to the present discussion is the contents of Hull's "Individual Differences" Postulate — Postulate XII. This postulate states something about how Hull would generalize his own theory. He suggests that although the general form of the relationships that he has plotted should remain the same when generalized, the *constants* in the equations would be expected to vary according to the species of the organism studied, the individual involved, and also with certain states of the individual so involved. It might be worth while to add to this list some of the entities of social psychology, i.e., the constants should vary with whatever "modal person" you are studying, or whatever "culture" or subculture the modal person comes from. At any rate, Hull's "vulnerability," which he greatly valued, exists with regard to the general form of the relationships, not the details. It is in this spirit that we have discussed the theory.

The problem of learnable drives. In the context of Brown and Farber's thinking we find it necessary, according to the requirements of the theoretical systems, to name the two relevant drives active in Johnny at the beginning of his action sequence. The problem of how to name such drives in the most useful way has been a recurrent one in social psychological thinking. Murray (1938) has provided us with some fruitful guesses on this score, and McClelland (1951) has extended some of Murray's thinking to an R-R operational level. The basic problem, however, of a full S-R analysis of such concepts as needs or motives is still in the developmental stage. Sears (1951) has pointed out that one of the fallacies in past motivational analysis has been the assumption of static traits or drives, whereas the problem probably should be stated in the context of *changing* structures, therefore in terms of a theory of learning.

Two present lines of thought and research in S-R theory may eventually help solve some of these problems. Within the framework of Hull's postulate III (1950), Miller (1951) has discussed how "fear" can be viewed as a learnable drive. Miller views fear as a response which is probably innately conditioned to painful or noxious stimulation. Insofar as it functions as a response, then it should follow the laws of response, one of which is that responses are conditionable to previously neutral stimuli. Miller also views fear as a response which produces strong stimulation, and therefore it has impelling or drive properties. Further, Miller views fear as a "cue," that is, as a provider of information which can also function as a generalizing or mediating stimulus. He has produced an impressive array of evidence to this effect, and this example of a learnable drive was found to be quite useful in Dollard and Miller's (1950) analysis of psychotherapeutic procedures from an S-R point of view. "Fear" becomes established as a response to previously neutral stimulation only under conditions where the fear is reduced as a result of some trial and error or instrumental activity of the organism in which it has been aroused. Such an extension of the principle of drive reduction has been criticized by Mowrer (1947 and 1950), and much research is currently being done in this area to clarify the factual and conceptual problems.

There are obviously other classes of learnable drives besides fear. Miller (1951) himself has provided a careful summary of all the empirical data with regard to motives which develop as the result of the operation of secondary reinforcements of other kinds than merely fear reduction. Probably no aspect of Hull's behavior theory needs greater clarification than the area of "secondary phenomena." In the study of motives which develop as a result of positive reinforcements from the social or intrapersonal environment, the present tendency is to infer the strength of the motive from the frequency or intensity of instrumental activities, which are classified together by reason of the fact that they have a similar effect on the environment. As an example of this, let us imagine our friend Johnny in a nursery playroom playing with dolls. We shall use the doll play situation as a "projective" test and will categorize as "aggressive" all the things which Johnny has the dolls do which have a destructive effect on the other dolls or the nursery schoolroom. We can use some index of the frequency of such acts as a definition of the strength of Johnny's "aggressive motive" in this situation. One of the theoretical problems will be to relate Johnny's behavior in this situation with the relative frequency of his aggressive acts in other situations (R-R analysis), but the major purpose of S-R analysis will be to ferret out the antecedent conditions that led to the current strength of aggression that Johnny displays in the present field. This will necessitate consideration of other emotional and "cognitive" factors which are interacting with drive-aggression in Johnny's personality structure. Several complex social motives are being approached in this fashion — dependency, guilt, affiliation, and achievement have been dealt with either on the response side or in terms of a full S-R analysis (Sears, 1951; Whiting and Child, 1953; McClelland, 1951).

MOWRER'S DUAL THEORY

A brief discussion of O. H. Mowrer's recent theory (1947) may provide us with closure for our section on Hull's theoretical system, since his dual theory (like that of Schlosberg, 1937, Skinner, 1938, and others) is intimately related to S-R reinforcement theory. Until 1947, Mowrer was one of the most active exponents of a monistic theory of the drive reduction variety. Since then, however, he has provided a clear and consistent exposition of a dual theory which can be succinctly described as an S-S *and* S-R theory. The theory has not been worked out in a formal fashion such as Hull's, but such formal expansion is implicit in his writings.

Two approaches will assist our understanding of Mowrer's thinking. First, we may note that Mowrer places great weight on the fact that in the Pavlovian and Thorndikian experiments two different classes of response are measured. If Johnny were a subject in the Pavlovian situation, we would be interested in the measurement of his salivary response (controlled by unstriped musculature). In the Thorndikian situation we would have our instruments prepared to record Johnny's walking, talking, and chewing behavior. These differences, and certain relationships that obtain regarding these kinds of response, lead Mowrer to make a fundamental distinction between *emotional learning* (Pavlovian conditioning of unstriped muscle responses, mediated by the automatic nervous system) and *problem solving* (Thorndikian learning of striped muscle responses mediated by the central nervous system).

Second, we must recall that Mowrer (1939) had earlier proposed that the emotion of "fear" be used as an intervening or mediating variable, the reduction of which would be a form of reinforcement. This concept provides an interpretation of many experimental facts (see Miller, 1950), and it has led to many experiments (see Mowrer, 1950). The problem that worried Mowrer, however, was this: What are the laws which will best predict the strength of learning of a "fear"? In the avoidance conditioning situation a conditional stimulus (e.g., a bell) is followed in time by a shock (US). Eventually the animal makes the motor response necessary to avoid the US, and this response occurs when the CS is sounded. The motor response leads to reduction of shock stimulation (and also, presumably, of "fear"). A consistent monistic drive-reduction theory would say that this drive reduction (i.e., reduction of "pain" and "fear") *also* led to the learning of the fear response to the CS.

It is here that Mowrer departs from drive-reduction theory. Mowrer (1950) argues that it is against common sense, as well as some experimental facts, to think that the fear is conditioned to the CS as a result of eventual reward. *Emotional* learning (or the learning of a "problem") is achieved neither by Guthrie's principle of S-R contiguity, nor by Hull's principle of drive reduction. Rather, it is learned solely by the contiguity in time of two stimuli, the CS and the US, and the contiguity laws worked out by Pavlov and others apply to this learning. Mowrer's thinking here is quite close to that of Tolman (1932), and we can use Tolman's common-sense terminology to make the notion more clear: the CS comes to function as a sign of the US through the mediation of the fear response, which provides the "meaning" to the sign.

But let us not leave the animal with *only* a problem of feeling afraid. He also sometimes learns to do something about it — he learns to jump when the CS is sounded and to escape the shock. This learning, Mowrer feels, cannot be predicted by an S-S theory, but requires principles of the Hullian type, leaning on reduction of the strong stimuli aroused by the fear response. The relevant laws are those which relate a response to a drive reduction — those laws which deal with response selection or problem solving. In short, S-S laws will predict the learning of our emotional problems, and S-R drive-reduction theory will predict the way in which solution of such problems will occur (when it does occur).

Another way of describing Mowrer's two kinds of learning is to call conditioning "cue-drive" learning, and problem solving "drive-response" learning. Mowrer has talked mainly of *fear* as the drive in "cue-drive" learning, so we must expand his present model to discuss Johnny's ice-cream seeking behavior. The problem is how to name the ice-cream appetite which motivates the sequence of responses. Presumably, Mowrer would consider the bond between the salivation response and the ice-cream bell as an example of "cue-drive" learning, the strength of which could be controlled (and predicted) by arranging the temporal relationship between the bell (CS) and arrival of ice cream on the tongue (US). The important point here is that eating the ice cream, or reducing the drive, is not the important factor, but is merely the *onset* of the US. This gives the ice-cream bell meaning or, in Mowrer's terms, conditions the drive to the cue. The distinction between the "onset of the US" (i.e., the arrival of ice cream on the tongue) and a "drive-reducing state of affairs" is more clear when we think of the avoidance situation discussed above, where the "onset of the US" is the application of shock, and the "drive-reducing state of affairs" is the arrival of the animal in a nonshocking, nonfearful place.

The overt responses of running, asking mother, and getting the nickel, as well as the behavior involved in paying the vendor, would all be learned according to the basic principles discussed by Hull. Mowrer does not, however, attempt to expand these principles to include "cue-drive" learning, as Hull does.

Mowrer also differs from Hull in his use of the important concept of secondary reinforcement. Hull uses a concept of secondary reinforcement which permits a previously neutral stimulus (like Johnny's nickel) to take on the capacity to trip off the reinforcing mechanism. That is, obtaining a nickel affects Johnny, for at least a time, as if he were eating ice cream (or reducing some other drive). Mowrer (1952), however, does not think that this occurs, but rather that secondary reinforcement should be given the meaning of "a reduction of a secondary drive." This requires, therefore, that whenever a secondary reinforcement occurs, some *secondary* drive must be present in order to be reduced. Mowrer (1952) suggests that whenever any drive is strongly aroused, *fear is always a concomitant,* that fear is a drive which is conditionable to external stimuli, and that it is also a stimulus which can subsequently be reduced. When it is reduced, the neutral stimulus objects (such as nickels) can become conditioned signs of a fear reduction. Their subsequent presence will thus maintain behavior in the way usually found in secondary reinforcement experiments.

A crucial test between these two approaches to the vital question of secondary reinforcement is very difficult to achieve, but could be of great importance to behavior theory, since it impinges on the area of greatest interest to social psychologists.

Mowrer has continued to emphasize fear as his major example of a secondary drive, but he often speaks of "secondary drives." This suggests one of the central problems of behavior theory and of social psychology: How can we approximate a more adequate classification of such secondary wants and how do such new wants become learned? Perhaps fear is the only secondary drive necessary, but the implications of such extreme parsimony have not yet been worked out.

Perhaps the clearest evidence regarding this theory of Mowrer's is the finding by Solomon and Wynne (1950) that when a measure is taken of the latency between the CS and the motor response as avoidance learning progresses, the curve is not a smooth single function, but shows two points of inflection. One of these may reflect the learning of fear, with the other reflecting the learning of the motor response to this fear. Solomon and Wynne find, however, that even when the autonomic nervous system is inactivated by use of drugs, some avoidance conditioning occurs. This may mean that what Mowrer calls emotional learning involves some skeletal responses as well as emotional ones. Alternative explanations of much of the data is still possible, however, and Miller has been able to make a spirited and ingenious defense of the monistic drive-reduction position (Miller, 1951b).

Mowrer's theory has generated considerable research, and the social psychological implications of the theory have been partly worked out by Dewey and Humber (1951), and most completely by Mowrer himself (1950). One might include Skinner's (1953) recent discussion of personality and social psychological issues as a treatise argued from much the same position as that of Mowrer, although Skinner's empirical notion of reinforcement and his attempt to avoid intervening mechanisms, particularly physiological ones, makes the discussion appear to be more different than it is.

GUTHRIE'S GENERAL THEORY

Guthrie has recently remarked that the difference between his own theory and that of the drive-reduction theorists "is of great ultimate importance but of slight practical and immediate importance" (1952, p. 235). From the point of view of systematic generalization to social phenomena, this does not follow, because Guthrie is here discussing the *essentials* of the two doctrines of reinforcement, not the *form* of the theories he has compared. One of the main values of a theory like Hull's is its heuristic value for those who wish to create a social psychology on the basis of funded experimental facts. Heuristic value in this case is a matter of great practical importance.

Guthrie, for example, agrees that his theory is not set up to handle felicitously some of the systematic changes in response phenomena which occur when the same stimulus situation is repeated again and again. Hull has focused upon such phenomena to build his system, whereas Guthrie agrees when Dollard and Miller (1951) simplify the system by dropping out such factors, and focus, as Guthrie does, on the single principle of the necessary and sufficient conditions for acquisition (1952, p. 225). Yet to the student of child development and socialization, the fact of repeated feedings of a baby by the same person at about the same time every day is a type of fact which permits a basis for ordering the empirical field. In this sense Hull's heuristic model is a very practical gadget. Guthrie can generate these phenomena, but only with confessed difficulty (1952, p. 229). In short, Guthrie has not developed a *predicting* theoretical system but prefers to use with great intuitive wisdom a basic principle, and to show with startling acuity how his principle can be viewed as operating in any learning situation.

Guthrie's writings do not fit the title "theoretical system" in the same sense as Hull's. His thinking might be characterized as a notion plus wisdom. We include a discussion of his theory here because there are various signs that his thinking is being built into one or more theoretical systems (Estes, 1950; Mosteller and Bush, 1951), and will probably become more *systematically* important in the future. The theory has already been important for some time in its effect on other theories. It has also been experimentally fruitful. Guthrie is somewhat of a latter-day psychologist's psychologist.

The "stimulus-response" formula. The problem of "stimulus" and "response" in Guthrie's thinking has remained closer to that of J. B. Watson (1913) than has that of any other current S-R theorist. Guthrie puts it this way when he is comparing his own thinking with that of Tolman:

> This book [his own] has undertaken a very different form of restriction for the purpose of understanding, namely, the restriction of the basic facts to stimuli to sense organs and to animal movements, with the limited admission of certain movement groups or classes defined by their end results. This is the highly controversial stimulus-response formula. Unlike Tolman's view, movements are not regarded as alternatives to the formation of likes and dislikes, to the perception of time and space, to the possession of the equivalent of a map, or to the sets which make short cuts and localization possible. Every horse owner knows that his animal has a strong tendency to maintain an orientation toward the stable where it is fed even when this place has been out of sight for some time. But orientation toward the stable is not something the horse uses instead of movement patterns. Men can demonstrate virtual possession of a "map in the head" by speech. But in this book speech is regarded as a form of motion. It is much more. But it is also motion. (Guthrie, 1952, p. 196. See Tolman, 1932, for comparison.)

Although probably all S-R theorists do accept the general notion of the S-R formula expressed in Guthrie's statement, there are differences, of course. Hull accepts it insofar as his associative concept (i.e., ${}_SH_R$) defines an association bond (H) between an afferent (S) event and an efferent (R) event. Hull has gone beyond this in other concepts, however. Drive (D) is defined (Hull, 1951) in the postulate in which it is introduced as a strong stimulus, but the stimulus is not *further* defined except by way of the environmental condition or operation of deprivation. There are further nonassociative elements in Hull's system. For example, I_R has no specific S subscript. Although Solomon (1948) has provided a peripheral stimulus definition for I_R, this is not the only possible interpretation of Hull's concept. Such nonassociative factors and such "operations" may refer to the fact that Hull is not a strict S-R theorist but rather an S-O-R (stimulus-*organism*-response) theorist in the sense that Woodworth has used these terms. (See Spence, 1952b.)

Further, the "S-R formula" above does not note the newer uses of the term "stimulus," such as that of Skinner (1938), who uses the term to refer to any part of the environment which, when modified, induces the occurrence of a given part of behavior. There has, therefore, been a de-emphasis of any particular existential definition of the term *stimulus* and of the term *response*. Possibly this freedom to try out different notions or functional definitions (see Miller, 1951b) of stimulus and response has freed S-R theorists for the broad lines of research they have explored in recent years. It is certainly a basis for the notion in this chapter of the possibility of fruitfully generalizing basic experimentally derived theories.

Guthrie himself reflects this change in part. He accepts the notion of "actions" as functionally equivalent and overlapping (ergo, "organized") sequences of movements. These actions occur on an all-or-none basis, but stable and predictable *repetition* of such sequence depends on the conditioning of a "key" or lead response to a broad number of the stimulus aspects of the situation. In this sense Guthrie deduces from his basic theory the possibility of a unit of complex instrumental responses, but only under conditions where such instrumentality is actually achieved. The explanatory level remains that of "movements."

Guthrie also reflects a change in the use of the term "stimulus." In his recent book he writes:

> There are two requirements which stimuli must satisfy in order to serve as scientific data. The first is that they must be physical changes affecting and activating sense organs. The second requirement serves a different purpose. In order to be usable as psychological data stimuli must be describable in terms acceptable to all scientific observers, whatever their theoretical position . . . [such as] the sound of the whistle . . . the sight of the pattern on the door . . . being recumbent. (1952, p. 276)

Guthrie sees the second criterion as one which serves a practical purpose in scientific communication, despite the fact that these examples of "stimulus" are already highly organized for the scientific observer as the result of past experience.

Guthrie, therefore, reflects to some extent the more "molar" uses of these terms S and R, but he still has as his major suggestion that in studying behavior change, we note very carefully *just what the animal does*. That is, saying, "Johnny went and got some ice cream," is not an adequate description of behavior for Guthrie. It refers merely to the fact that Johnny was *successful*, but it does not tell enough about what he *did*. When Skinner (1938) records how many times a rat presses a bar, he is counting "environmental events," or the effects of the rat's movements on the environment. But Guthrie argues that the rat could have achieved the change in the environment by use of his paw, his snout, his ear, or one of many other "movements." That is, calling the response a "bar press" or an "ice-cream getting" covers up the important event, that is, whether the response was "pawing at the bar," "a snout-press," "an ice-cream snatch," or "an ice-cream run." Imagine how far we are from Guthrie's type of objective description if our protocol merely says "Johnny stole an ice-cream cone"!

Guthrie agrees, of course, that stealing and ice-cream-getting occur, but his point is that these should be given a special term — "actions." He also argues that, in learning, sometimes it is actions which are learned, that is, whole patterns of movements (which carry with them their own stimuli to which the next movement is conditioned) are sometimes conditioned to a new stimulus, and we might refer to an entire bundle of responses of this kind in terms of its results.

The principle of S-R contiguity. But the most important thing to keep in mind, from Guthrie's point of view, is the deceptively simple principle that learning occurs when (and only when) a stimulus occurs in close contiguity with a movement. We have already outlined this principle, which Guthrie uses to explain the empirical facts of the acquisition and extinction of a behavior sequence. Let us consider it more carefully in the present context.

Guthrie states his principle of learning by contiguity of S and R in the following way:

> The present theory assumes that the repertoire of responses is set by the physical structure of the

> responding organism and the indefinitely large number of combined patterns of which the effectors are capable. It assumes also that once any pattern of response has been evoked in a stress situation (all stimuli set up stress) which includes a given signal or cue, the given signal or cue becomes an elicitor of the response pattern and remains an elicitor until the signal has been presented with a new response. Acts and movements must terminate and make their continuance impossible. One can descend the stairs but cannot immediately repeat it because he is now at the bottom, not the top. The raised arm cannot be raised; it can only be lowered.
>
> In any general situation, therefore, there will be a series of responses, and each new response will alienate its predecessors from their association with any signal present when the new response is given. In stress situation S_m (m indicates the stimuli here called "maintaining stimuli"), a series of responses, R_1, R_2, R_3, . . ., R_s, R_t is evoked. If R_s removes S_m or removes the animal from S_m, then S_m remains a signal for R_s, its last association. (1952, p. 297)

Guthrie's analysis of other learning phenomena. After we presented our example of Johnny's sequence of behavior we outlined a series of empirical phenomena of learning which Hull, we saw, considered important enough to be included as separate postulates in his system. Guthrie recognizes these phenomena, but he does not consider them as sufficiently representative of the basic behavior law he is working out to be considered as essential reference data. Stimulus generalization, so important in the Hullian system, is dealt with by Guthrie in a different manner. The basic empirical fact is that other stimuli which are similar to the conditioned stimulus will evoke the same response, to the degree to which such stimuli are similar. Johnny, for example, would probably salivate to a knife-grinder's bell unless he had specifically learned to discriminate between the two sounds. If this occurred, Guthrie would describe the phenomenon as due to the fact that there is a common response which has occurred to both bells, or at least responses which have identical components. That is, Johnny tends to "listen" as a response to both sounds, and this orienting (or "key") response is chained to the conditioned response through the mediation of "response-produced cues." In short, the orienting response of listening leads to the evocation of the same series of responses, including the conditioned response, but the causal agent is the similarity in the orienting responses, not in the similarity of the external sounds themselves. If Johnny had made a discrimination between the two stimuli, then there would be a difference in the orienting responses such that the generalization phenomenon would not occur. Johnny might say "knife grinder" to S_1 and "ice-cream man" to S_2, while listening.

We shall see below that Osgood (1952) has attempted to set up ways of measuring independently these mediating responses, which in an S-R theory (Hull's included) are a prototype of "meaning." It should be noted, however, that Guthrie's theory does not accept generalization as a special process. He tries to show how generalization phenomena, when analyzed carefully, are not exceptions to the rule that we always respond to the *same* stimulus with the *same* response. The problem is to keep our eye on the changes that occur in responses and in stimuli. Generalization appears to imply that organisms sometimes respond with a conditioned response to a stimulus to which the response had never been conditioned. To Guthrie this interpretation is only an appearance — there is always a specific mediating stimulus which sets off the observed response. In short, transfer of training is always accomplished by identical elements. The illusion arises when we record the wrong stimulus or the wrong response or have not noted that the mediating response is conditioned to many "external" stimuli, or that the stimulus produced by the mediating response is conditioned to many other measurable responses.

The reader who is unacquainted with Guthrie's thinking might wonder how Guthrie would explain such phenomena as the apparent stabilization of learning which occurs as the number of trials increases. It should be remembered that Guthrie sees reinforcement as an all-or-none phenomenon. One might expect him to agree with discontinuity theorists of learning that learning has occurred when the cognitive (stimulus) field is suddenly restructured. To think this would be incorrect, since Guthrie talks about the fact that the learning curves we usually find are due to the fact that the response being learned is becoming conditioned to more and more of the stimulus components of the situation, and therefore the performance seems to be improving slowly. This means, then, that in most situations, learning, although it occurs in an all-or-none fashion, appears to be cumulative, unless we can watch more carefully for the movements and stimuli which are suddenly and completely conditioned or the useless movements which drop out for one of the reasons we have given above. This constant emphasis on the necessity for behavior analysts to look carefully at what the animal last did in the situation is the major advice Guthrie has to offer.

Guthrie's primary reference data (Guthrie and Horton, 1946) are contained in a series of experiments in which he and Horton photographed the behavior of cats who, upon being placed in a box, learned to press a pole away from the vertical and thereby escape from the box. Guthrie found that his best chance of predicting the exact pattern of movements the cat would use on the next trial was to predict that the cat would do what he last did. This prediction would not always work, of course, but it worked very often. Further, Guthrie found that the cats tended to repeat over and over a limited set of movement sequences, so that knowledge of these possible sequences might assist the prediction even more. Note that Guthrie, in studying cats in the puzzle box, was interested not in the speed with which the cats would "succeed," but mainly in studying the *exact sequence of movements* which the cat used in achieving success. Perhaps the increased use of moving pictures in social psychology and anthropology will render Guthrie's thinking more directly relevant to analysis of problems involved in teaching new skills to members of cultures with highly stylized movement patterns, to uncovering laws of the "language of movements," which has hardly been touched upon except in the limited field of verbal movements, or to uncovering the antecedent laws related to Allport and Vernon's (1933) basic factors of expressive movement.

Guthrie's studies of cats emphasize the richness of the cat's movements and also the tremendous degree of stereotypy which characterizes behavior when viewed in this careful manner. Other research which his theory has generated includes attempts to test Hull's notion that reinforcement is a necessary condition of learning (Sheffield, 1950, 1951), studies of methods of response extinction which other theories would not generate (Seward, 1949), analyses of the role of response factors in the recall of uncompleted tasks (Boguslavsky, 1951), and the explanation of Thorndike's "spread of effect" by invoking the notion of "guessing habits" (Jenkins and Sheffield, 1946), to name only a few. Boguslavsky (1953) has recently developed a predicting probability model for use with Guthrie's notion of orienting response. Guthrie and his colleague Stevenson Smith have used the notions of the theory in discussing the concept of personality (Guthrie, 1944), the problems of psychological conflict (Guthrie, 1937), and a very sane discussion of the problem of eneuresis (Smith, 1948). Some of this writing, as Hilgard (1948) has suggested, bears no direct or logically necessary relation to Guthrie's theory, but it shows that the principle of S-R contiguity can function very comfortably in the personality analyst's kit of tools.

This method of generalizing is usually an advance over common sense, or at least casts a fresh light on common-sense answers to questions. This is possibly due to the fact that S-R theorists, either of the contiguity or the reinforcement school, view such problems as "seeing," "wanting," and "purposing" as activities of the organism which *require* explanation, whereas many writers in psychology assume these activities as givens or as properties which all human beings display (which, of course, they *do*, much of the time) and use such activities as explanations rather than as phenomena to be explained. Until social psychology has organized its data in a systematic and reliable fashion, the *systematic* extension of general behavior theory must give way to the use of "notions plus wisdom." Further, Guthrie's long-continued emphasis on response-produced stimulation has been used to great experimental and theoretical advantage by other psychologists, particularly by Miller (1951b). Guthrie has himself outlined the relation of this emphasis to the more recent notion of "feed-back" as developed by Wiener (Guthrie, 1952, p. 296).

We began our discussion of Guthrie's thinking by criticizing *his* over-all comparison of his theory with that of Hull. We shall reiterate this criticism and then go on to agree with what Guthrie probably meant. Guthrie deals mainly with a basic principle whereby behavior change occurs, whereas Hull attempts to deal not only with such a principle, but also with whatever other processes and variables must be considered in order to arrive at a rough first approximation of a theory which will *predict*, and possibly control, such changes in behavior. Guthrie remains parsimonious, and perhaps eventually correct, on the issue of whether or not drive reduction is "necessary" in a basic law of "acquisition." Hull drops the parsimony but provides us with a gadget with which social psychologists can play at the "game" of predictive science.

But Guthrie's notions are, in another practical sense, close to Hull's. Guthrie feels that Dollard and Miller (1951) are essentially talking his language, for all practical purposes. The reduction of a drive is one way of changing the situation so as to preserve the S-R connection for the next trial, particularly if "fear" is the drive or maintaining stimulus which is reduced or altered. Further, the S-R theorist in the psychotherapeutic situation such as discussed by Dollard and Miller (1951) has little use for the quantified laws relating antecedent conditions to intervening variables, although the serious student of personality development or socialization may find them of considerable use.

More explicit attempts at pulling the S-R contiguity position closer to that of the S-R reinforcement theory are under way. Mosteller and Bush (1951), in a display of theoretical fortitude, have attempted to partial out the strengths and weaknesses of the two theories and to put the virtues together in the format of a probability model close to that of Estes (1950). The future may see an amalgamation of the S-R contiguity notion added to a reinforcement notion, so that one could say that the probability of the response occurring on the next trial is greater if it (1) occurs and (2) is also "reinforced," than if it only *occurs*. When other processes are explicitly added to this kind of theory, the question of the correct theory of reinforcement remains a matter for quantitative experimental study at the level of the theoretical system, and a more useful set of equations is required for purposes of generalizing (i.e., of constructing what we have called theoretical-empirical systems).

The Empirical-Theoretical Systems

THE PROBLEM OF GENERALIZING IN SOCIAL SCIENCE

We have outlined in the context of a sample of everyday behavior the theoretical notions of three important S-R theories. We have noted that the data for much of these theories have been induced and that the data according to which they are formally defined are limited. One of the next steps to be taken by those interested in such theories would be to deduce and test further consequences of the theoretical position and thereby increase or decrease the confidence one has in the applicability of the model to the defining data. This endeavor is going forward in many laboratories of the U.S., as well as elsewhere in the world, and the theories discussed have, over a period of years, been changed as a result. A second step which one can take is to work out systematically the relevance of the theory to new classes of data, that is, to classes of data which are currently viewed as being different in some respects from the data used in devising the theoretical system itself. It is to this endeavor that many social psychologists have turned in their attempts to advance our systematic understanding of problems and areas of data usually discussed under personality and social psychology. This is the problem of generalizing theory in its most generic and systematic, and therefore most simple, sense.

This problem is not one which is limited to S-R theory, of course. It is a problem faced by anyone interested in the advancement of a systematic theory in the field of social psychology. The problem of "relevance" of a theory to a particular set of observations, or of a particular set of observations to a particular theory, can be a most distressing one. Anyone working with a certain social or psychological problem will develop a language for discussion of the matter which appears to him to be the most relevant, and which becomes for him *the* language for describing the problem or area under investigation. Such descriptive language may be developed for a variety of reasons, and these reasons need not include the attempt to explain systematically, but merely may reflect an interest in capturing what meaning one can from the observations he has made.

Many formulations of social psychological problems have been made in this descriptively relevant fashion, and many of the problems of social psychology remain in this state of limited usefulness. Much of the theory which has grown from the study of social phenomena in the "natural" setting still suffers from the fact that it is a happenstance result of the descriptive method, rather than having been systematically constructed on a general basis. Such theories present, however, a sense of immediate relevance, and diverging from them may give one a sense of reducing the richness and complexity of the original problem to something less relevant. Many of the theories of this kind are based upon what S-R theorists, from their own position, describe broadly as R-R (response-response) relationships. One may, for example, have studied the problem of intelligent behavior and have developed measures of an aspect of this behavior, working out a series of relationships among such measures. This immediately practical endeavor provides the researcher with speedy and amply deserved reinforcement, but from an S-R point of view such formulations often reflect a methodological limitation of the field under study. Very complex analyses emerge from the structural approaches which R-R theory develops by using these measurements derived from study of the relatively uncontrolled field. Some theories, such as Lewin's and the structural aspects of Freud, as well as the empirical research areas of intelligence testing and projective testing, have in common the limitations imposed by the decision to remain within what one might call "field" or "natural observation" methodology.

Whereas attempts at descriptive pictures of already-produced structures have an apparent, and often genuine, relevance as their major virtue, they are too often at the present time limited in the extent and the precision with which they can predict more than momentary changes in the structures so set up. S-R theorists have tended to take another approach to the problem of theory construction and therefore the problem in generalizing is a somewhat different one. Recognizing that laws of changes of structure may require a different unit of analysis and different structure of theory from that of the attempt at description of the structures themselves, they have placed primary effort on the experimental study of situations where such change in response hierarchy or response probability formulations occur. To maximize the possibility of experimental manipulation of variables and controlling of irrelevant factors, they have been forced away from the maximization of immediate relevance of their units and their language. Spence (1952) has noted that this decision was a relatively self-conscious one, and in light of some of the current advantages that accrue from S-R theorists' reanalysis of structures, this decision does not need defense.

Having found the relationships among some, at least, of the important variables involved in those aspects of behavior change which are more or less environmentally manipulable, the tendency recently has been to try out, either by experimental means or by naturalistic observation, new classes of referents in terms of their interrelations. This is the problem of explicit generalizing which marks, in the present writer's view, much of the endeavor of social psychologists interested in S-R theory. Ceding value to this approach does not in the least negate the reasonableness of the sophisticated attempts at structural analysis of such people as Lewin, Parsons, Malinowski, Thurstone, or Tolman. To study structures first and their change second is quite as reasonable as to study changes of structure first and to refine the notions of structures later, and there are many times when one kind of analysis can complement the other. However, Bergmann and Spence (1944) may be correct in their assertion that, in principle, if you have the relation between S_1 and R_1 and also the relation between S_1 and R_2, you can often solve for the relation between R_1 and R_2, whereas knowledge of the relation between R_1 and R_2 does not solve the S-R problem. And perhaps, for example, the failure of Pareto's attempt at a mathematical theory for sociology was due to the impossibility of dealing with the large number of variables which structural analysis often entails. Perhaps the painstaking emphasis on experimental criteria for the acceptance of variables and the statement of their relative importance will save S-R theory from some of the confusing problems that other attempts have met. In this sense, one of the major values to social psychology of the experimental method is that it assists in reducing a set of variables to a usable number in attempting a first approximation to a general theory of behavior.

It has been fairly widely accepted by S-R theorists, or at least by the reinforcement variety of S-R theorist, that a hypothetical construct or intervening variable must eventually be defined in terms of both antecedents and consequents. This, however, does not restrain preliminary analyses of a new social or personality problem area from being defined tentatively "on one side" or the other. Thus it is possible for an S-R theorist to define on the response side, in whatever complexity he finds best, a new (for S-R theory) concept such as "drive identification" and to search, with the aid of his theoretical system, for the particular pattern of antecedent variables necessary to generate the sample of complex patterns of responses he has tentatively defined as "drive identification." One may view Freudian theory as having emphasized antecedent definition to the de-emphasis of both antecedent and response-side operational specification, and such current forms of response-response theory as Lewin's as having emphasized response-side definition to the exclusion of those laws which would generate some of the possible patterns of antecedents for a given "momentary field." Tolman (1951) has seen this problem and has apparently attempted to construct the general outlines of a model to include a solution of all these problems. However, his insightful cognizance of the full range of problems, and the aim for comprehensiveness which his thinking displays, probably cannot continue too long to substitute for a precise statement of relationships among constructs.

The above propositions may be summed up by stating that in selecting his language of description or of explanation, the social psychologist can usefully recognize that structure analysis without process is only part of our task, and that process analysis without eventual refinement of structural description is not fully satisfactory either. Further, the decision as to choice of descriptive and explanatory language should not be too colored by what might be called a form of "aesthetic fallacy." That is, theoretical analysis does not require that the phenomena under study be "phenomenologically" reflected in the theory. Most useful explanations go beyond full and aesthetically satisfying description, and Marx (1951), among others, has described the scientific orientation as one at a polar extreme from the aesthetic or literary.

Whether or not these preceding remarks are accepted, there will be agreement, the writer feels, on the proposition that a theory which is anchored to a set of defining data must face the problem of checking the legitimacy of any translation of this formulation to a new area of data. Let us look then at this basic question for a moment: How can one generalize, or in other words, how can one construct what we have called an empirical-theoretical system?

SOME WAYS OF GENERALIZING

Operational extension. The first and most important way of generalizing is accomplished, as the expression used above suggests, by extending the class of operations of the theoretical system. That is, methods which are similar to those of the defining operations of the theoretical system are utilized in situations concerning which the stated principles or relationships of the theory *seem* to make predictions. A simple example of this is the following: the theoretical system may say, for instance, that organisms in a state of hunger drive move toward food more rapidly than those which are not in such state. The defining operations of the theory, to simplify, may be those concerning rats in a runway situation. Now we may extend this to predict that children coming home from school at noon, when they are hungry for their lunch, will make the trip more rapidly than if they come home from school at some other time when the hunger drive is lower. Or, to consider Johnny's situation, if the ice-cream man comes around just before lunch or dinner, he may run faster to his mother for the nickel, and on to the ice-cream man, than if the vendor arrives just after he has finished a meal. The operations for measurement in these latter situations are very similar to those of the rat runway problem; they involve the measuring of comparative time intervals. The extension may be considerably less obvious than this one, of course. For example, having found such a temporal relationship in a rat runway problem with such an antecedent index as strength of drive, one could also, with human subjects, use latency of response in an association test which included food words and other words, having varied the degree of food deprivation of these subjects as the antecedent variable. One may even attempt to take some complete theoretical model which is defined in terms of specific reference data and apply it to a different species of organism, or to another aspect of the behavior of the same species.

If the theoretical relationship is as predicted, and particularly if additional deduced consequences are also upheld, one can feel relatively confident that the original theoretical model applies to the new area of data. This procedure is an exact and a difficult one, however, particularly for social psychology. Among other problems is the decision as to what operations are "similar" to the defining ones of the theoretical system. A program of research undertaken by Solomon and his colleagues can be placed in this category of operational extension and will be discussed further below. Another example is the study by Dinsmoor (1952), who extended the operational denotations of some of the "new look" perception findings to a Skinner-type conditioning situation, obtaining similar relationships to those found in the former situation.

A major advantage of this method of generalizing is that one can feel relatively sure that the theoretical model does not apply to the new area, or is at fault in some way, if a clear malprediction occurs. It may be useful to spot check in this way in some new area before going on to less conclusive, although no doubt easier, methods of generalizing.

Descriptive analogy. This type of generalizing is at first glance appealing because of its simplicity, but too often the simplicity is deceptive. It has been widely used, however, and certainly merits discussion. Here one finds that a description of a theoretical concept, a bit of defining data, or a relationship sounds similar to the description of some other concept, bit of data, or presumed relationship outside the theory. One then, by such analogy of description, adds to the latter some of what is known to be the case with the former, and assumes generalization to have taken place. The extratheoretical data and the theoretical propositions often are found to fit quite well, but here considerable caution is needed, since the similarity may be only that which first suggested the analogy, with nothing having been added — except a heuristic device of doubtful propriety. Many examples of this method are found in discussions of the social structure in analogy to the human organism. F. H. Allport (1942) has pointed out problems raised in this sort of proceeding. A very early instance of this in modern social thinking is Hobbes' comparison of the state with the human organism. Some other "simple and sovereign" concepts of social science are of this nature. One takes a notion, such as power, sympathy, or imitation, from a situation where it is applicable and applies it to other similar-appearing types of relationships as an explanatory device.

To use this method of descriptive analogy at all usefully, some restrictions should be observed. To make predictions which are more precise than the early observed similarity which prompted the analogy, the theoretical context from which the analogy-term is drawn should contain a set of at least two variables, or one quantifiable variable. It may be that the original variable has implicit properties which will fit it for such use, if these properties are first explicated carefully. It probably goes without saying that where such quantification, or differential prediction, is possible it should be undertaken, if descriptive analogy is to be useful.

Operational analogy. In this method the analogy is between the measure we would like to have and the measure which it is possible, or most feasible, to obtain. This may sometimes be a measure of approximation, where a part of a realm of data stands for the entire realm of data desired, but often it involves operations purely analogous to the desired operations. The method of operational analogy is a very reasonable compromise procedure for many areas of investigations. Thinking in terms of Brown and Farber's outline of Hull's theoretical system, we might, in doing a study of the socialization of aggression in the Hopi culture, for instance, desire a measurement of the frequency with which the child has been punished for naughtiness. Let us assume that it is not feasible to obtain an actual frequency count of such punishments in a random sample of Hopi children, the time sampling technique extending over a period of years. We may compromise with an analogous operation. We attempt with interview methods, or some other device, to ask certain strategic questions of the parents of these children, and, with some tentativeness, accept the estimates of the number of punishments for naughtiness as sufficiently analogous to the kind of measure we really would like to obtain.

An example of an excellent study using operational analogy is Pauline Sears' (1940) investigation of level of aspiration. Her operational analogy was in utilizing the records of school children's marks as standing for their full history of failures and successes, and she related these to patterns of level of aspiration determined experimentally.

Operational analogy meets the criteria suggested above for the reasonable use of descriptive analogy, and goes somewhat further toward the kind of integration of data and theory that many social psychologists desire. One of the ways in which the researcher can check whether or not he has achieved a reasonable analogy to the measure he would like to have, would be to see whether such analogous measures are related to one another, or to actual measures, in the manner suggested by the theoretical system. Thus one value of the theoretical system, as discussed previously, is that it can provide us with specific notions which state the relationships among a set of potentially useful variables, for broader operational approximation. A second value here of the theoretical system is that it can provide us with a fairly large set of explicitly interrelated variables so that any temporary success we have with our analogous measures in approximating the experimentally found relationships of the theory can be further tested by relating our measures to additional approximations, as suggested further by the theory.

At this point a problem arises concerning which S-R theorists have disagreed among themselves. Some argue that the S-R position rules out the use of human yardsticks in research. Thus Bergmann and Spence (1944) would argue that the use of human observers or interviewers in making judgments or ratings of behavior is a measurement which is one step away from the kind of direct measurement of behavior which they recommend. Other S-R psychologists feel that rating techniques and content analysis methods, such as the ones being developed for use with the TAT, are acceptable either for basic experimental purposes or for operational analogy. It is interesting to note that some quite influential research (e.g., Eleanor Gibson, 1942) in the recent human learning laboratory has found it desirable to utilize human judges in order to calibrate the degree of similarity among stimulus materials used in the experiment. It is amusing to note that Schlosberg and Solomon (1943) utilized psychological methods with human beings in order to calibrate degrees of gray on stimulus cards later used with rats, with predictable results. Perhaps the resolution of this kind of methodological controversy among social psychologists who have found S-R psychology stimulating would be to consider the psychophysical problems involved in the training of judges and one-way screen observers as the basic procedural research area it undoubtedly is. Most social psychologists consider such problems as merely additional operations, to be tested for reliability and then added to the operational limitations for generalization of the findings obtained with such methods.

Quandaries in generalizing. Vincent and Helen Nowlis (1952) have recently given a good summary of some of the problems which arise when one is faced with the failure of a particular hypothesis to be upheld when generalizing from a basic experimental system. They see at least three possibilities: (1) There has been inadequate translation of the concept or set of concepts from the simple laboratory situation to social situations, (2) the research may have failed to develop adequate instruments for indexing the very difficult phenomena being studied, or (3) the law which one is attempting to apply is not fitted for application in such complex situations. All these possibilities are to be considered when such failure to uphold the hypothesis occurs.

Sears (1951) has given an example of the utilization that can sometimes be made of data which fails to verify a hypothesis. To be brief, he found that this failure in confirming the hypothesis, when considered in conjunction with a suggested change in the basic theoretical system, motivated not only the search for further evidence in the area which he was studying, but also suggested experiments to be done at the level of the theoretical system itself.

SOME GENERALIZED S-R MODELS

It remains for us to outline some of the uses which have been made of S-R theories in attempts to increase our understanding of social psychological phenomena. We shall briefly discuss a series, by no means comprehensive, of such attempts, ranging from extensions of whole areas of theory to extensions which involve a very limited set of relationships.

A basic methodological factor should be mentioned at this point. Sophisticated use of S-R theory makes no assumption as to whether one must begin the analysis on the response side or on the stimulus side. Thus, most of the examples which we shall use reflect the acceptance by the S-R theorist of response-side definitions of phenomena as they have been discussed by social psychologists who have often been of different theoretical persuasions. Learned motives, attitudes, personalities, roles, neuroses, identifications, meanings, and clinical insights are entities which represent phenomena in varying degrees of clarification and measurement. When a psychologist interested in the extension of *any* theory decides to grapple with a problem in one of these areas, he must come to terms with the situation that the phenomenon defined may be of a most complex nature relative to any current statement of the theoretical system. We have noted in discussing Hull's theory that Brown and Farber found it necessary to suggest elaboration of the theory in order to make it relevant to certain emotional phenomena. Sometimes such extension is not necessary, but the investigator must often tread the careful path of deduction for a considerable distance in order to find a proposition which is consistent with theory and which bears upon the empirical phenomenon which he wishes to understand. Further, the extension sometimes deals only with certain aspects of the complicated phenomenon under analysis.

Basic definition for a social psychological research endeavor can therefore occur on the response side or on the stimulus side. One of the hallmarks of S-R theory, however, is that the analysis must eventually include a careful definition of the consequent phenomena under discussion as well as an explication of those antecedent factors which can relate to the consequent phenomena in a fashion relevant to the intervening variable set. In our sample of theoretical extensions we shall find that some of the researchers have barely started a consideration of the antecedent factors, as they have found the task of systematizing and validating response-side measures to be a major research task in itself.

Imitative behavior. Probably the most influential recent extension of S-R analysis to social behavior has been the work of Miller and Dollard in their analysis of the ancient social psychological concept of imitation. Their book, *Social Learning and Imitation,* was an attempt to redefine the place of this concept in social psychology. Whereas Tarde had said, "All society is imitation," Miller and Dollard overstate their purpose as trying to show that "All society is learning." An analysis of the kinds of behaviors which occurred when "imitation" was supposedly operating led them to a trichotomy of *copying, same behavior,* and *matched-dependent behavior.* They view these, in effect, as response-side phenomena derivable in terms of a drive-reduction theory of learning. Imitation, therefore, comes out of the analysis having shifted from its place (for Tarde) as an intervening variable to a position as a response pattern which occurs under given conditions as an accomplishment of the learning process. They do outline, however, the conditions under which a learned *motive* to imitate may develop in children and therefore be utilized for some analytic purposes as an intervening variable.

Matched-dependent behavior develops in a certain kind of interaction pattern, usually where a less skilled, or subordinate, person is in interaction with a person of higher status or greater skill. The subordinate person finds himself gaining rewards by doing what the superordinate person does, but he learns to respond only to the *former's acts,* and not to the *cues* to which the former is responding. Sufficient repetition and reinforcement of this behavior in such social contexts can lead to the development of a learned drive to imitate. This formulation, with changes, has only recently begun to generate attempts to experimentally induce learned motives of this class in the behavior of children, and to generate predictions in the behavior of children where experimental controls are not possible. (See Sears, 1950.) A recent study by Olds (1952) on the experimental control of learned drives in children is in part a resultant of this general conception of Miller and Dollard.

Miller and Dollard's theory of the development of matched-dependent imitation will probably be found relevant to the important task of developing an adequate theory for stating the social psychological conditions under which differing strengths of dependency motives are developed and in explaining individual differences which are found clinically in this area. The development of standardized instruments for the measurement of this motive are still under way.

Psychotherapy. More recently Dollard and Miller (1951) have suggested certain syntheses of behavior theory and psychoanalytic theory, following in the tradition set by Miller (1944), Sears (1943, 1944), Mowrer (1950), and Dollard and Doob, *et al.* (1939). They attempt to derive from S-R theory certain practices which have developed in therapy. Evaluation of such an attempt can be based on at least two criteria: (1) whether or not they have been able to place certain interaction procedures derived for limited use in therapy within the context of a much more broad and explicit theory, and (2) whether or not the translation leads to further insight in clinical interaction. (A successful translation of Freud into some form of modern psychological language could not fail to be of advantage, at least in the sense that it would reduce the authoritarian atmosphere which so often surrounds interpretation of Freud's thought.) These evaluations cannot as yet be made with any finality, but there is little doubt that these formulations have led to fruitful experimentation and to further theorizing.

This last statement can be briefly documented by reference to an empirical study and a theoretical exercise. Clark (1952) has recently aided in the task of untangling some of the factors involved in the complicated phenomena of guilt expression and the function of guilt. Using the theory of conflict of Dollard and Miller, Clark predicted that sexual drives would be more reflected in TAT protocols when the subjects were under the influence of alcohol. He further predicted that with those exposed to sexually stimulating pictures, but without alcohol, guilt would tend to reduce symbolic sexual expression *below* that found in a control group who were not exposed to sexually arousing stimuli. Both predictions were upheld, and the theory made possible a clarification of most of the data.

S-R conflict theory has been long in development, the earlier work having been summarized by Miller in 1942. Miller's more recent work (1948) has been focused on generating, among other things, some of the phenomena subsumed under the psychoanalytic concept of displacement. Displacement phenomena are related to the process of generalization, and the theory involves making certain explicit assumptions about how approach and avoidance responses generalize from one stimulus situation to other similar situations when the approach response is inhibited by a conflicting response. As an example of this important kind of small model construction within a broader theory, let us look at Miller's assumptions:

> (1) that the direct response to the original stimulus generalizes to other similar stimuli, with the amount of generalization becoming smaller the less similar the stimuli; (2) that the response which conflicts with the occurrence of the direct response to the original stimulus also generalizes to other similar stimuli, becoming weaker the less similar the stimuli; (3) that the gradient of generalization of the conflicting response falls off more steeply with dissimilarity than does that of the original response which it inhibits; (4) that when two or more incompatible responses are simultaneously excited, the one with the greatest net strength will be the one which will occur; and (5) that the net strength of a response will be its strength minus that of any response incompatible with it which is excited at the same time. (Miller, 1948, pp. 167-168)

With these assumptions Dollard and Miller are able to generate many practices and predictions, from the tendency among therapists to focus *first* on reducing a patient's anxiety, to predicting the major findings of Clark's study. Others have made use of this "model" also. It has been used in analyzing data from children's doll play (Hollenberg and Sperry, 1951), data from a cross-cultural study of myths (Wright, 1952), data from a cross-cultural study of theories of disease (Whiting and Child, 1953) and data gathered from the behavior of pigeons (Brush, Bush, Jenkins, John, and Whiting, 1952).

The above tests of the usefulness of the Miller assumptions have led also to further theorizing. One suggested direction of change is the work of Bush and Whiting (1953), which subsumes many of the suggestions of Brown and Farber which we have outlined above. Bush and Whiting have broadened Miller's formulation, taken out certain "flaws," and have elegantly ensconced their notions in an extension of a "set-theoretic model" developed by Bush and Mosteller (1951). In terms of several criteria, this theory ranks as one of the most developed of its type in social science: it has generated and subsumed several classes of data within an explicit framework of assumptions, is open to alteration, and also makes use of a probability language which is possibly the kind of language best fitted to social and psychological theorizing. Future experimentation will reveal whether this trend continues to be valuable, but in the meanwhile, Bush and Whiting state that:

> Any formal quantitative model of a complex psychological process necessarily involves numerous abstractions and simplifying assumptions and as a result is inadequate for describing many, indeed most, instances of that process.
>
> The price of precision is the loss of generality, and many psychologists may feel that the price is often too great. The argument hinges upon, it seems to us, the strategy one uses in his research, i.e., whether precision or generality has priority. There is always hope of making a general theory more precise and of making precise theory more general, but which is easier is a matter of faith not fact. Our faith, and apparently Miller's also, is in the procedure of slowly making more general a model with built-in precision. When simplifying assumptions are made, attention is focused on the factors which make the phenomena complex. Generality is increased by examining these assumptions, one by one, and finding suitable substitutes. Such a mechanism for improvement is hard to find for increasing the precision of a general theory. (Bush and Whiting, 1953, p. 261)

Extensions to child development. One of the most extensive attempts at systematic extension of S-R theory to the development of personality in children has been undertaken by Sears and his co-workers at the University of Iowa and at Harvard University. Although only a small portion of this research has yet been published, it represents a broad attempt to develop instruments for the measurement of certain personality properties of children and also the measurement of certain important patterns of antecedent variables which are suggested by the theories of Hull and of Miller and Dollard. Sears has reported the results of a portion of this study in his recent presidential address to the American Psychological Association (Sears, 1951). He pointed out that a very low positive correlation was found between the antecedent degree of frustration of middle-class children during socialization and the consequent frequency of aggressive activities shown in their doll-play stories. When, however, the antecedent variable of punishment for naughtiness is taken into consideration, considerable increase in correlation with the aggression indices is found. Doll-play aggression frequency appears to increase as a positive function of previous punishment for aggressive behavior, as well as the degree of frustration imposed by the structure of family interaction.

One possibility for explaining these two factors is by integrating them in a conflict theory like that of Bush and Whiting (1953) or of that of Brown and Farber (1951). This has been suggested by Hollenberg and Sperry (1951). As a result of these researches, Sears has been led to emphasize the importance of the different social conditions under which one would expect a learned motive, such as aggression, to be expressed in behavior. He found, for example, that both the high- and low-punished groups of children showed less incidence of aggressive behavior in the school situation than did those who were punished for aggression to a medium degree in the home. The problem, therefore, for S-R theory becomes one of developing measures and laws for predicting the transfer of a set of behaviors learned in one situation (e.g., the home) to other social situations such as school and doll-play situations. After outlining an S-R theory of dyadic interaction, Sears suggests:

> With a conceptualization of the dyadic variables involved . . . it is possible to make a statement that goes beyond the narrow confines of these measuring situations. In this instance, the more general statements can be made that, first, the amount of aggression will be a relative function of severity of punishment; and second, with severity of punishment held constant, the amount of aggression will vary positively with dissimilarity of the dyadic situation to the punishment situation in the home. (1951, p. 476)

This raises a problem which we shall discuss below and which Doob (1952) has recently discussed at some length, the problem of developing a rationale and adequate measures for the "meanings" which different kinds of social situations have for individuals.

The research of Sears and co-workers has dealt with other phenomena of social psychology, as these phenomena appear in the socialization of children. In company with Mowrer (1950) these researchers have found it necessary to focus attention on the mechanism of "identification" as a means of explaining the learning of certain personality properties and role behaviors by children where there has been no direct reinforcement for behaving in these ways. Mowrer has suggested that the sounds and sights of parents' behavior take on secondary reinforcing value and therefore when a child indulges in these behaviors he is self-reinforced, and certain learning is thereby accomplished. In short, identification may well become an additional learned drive in S-R reinforcement theory and, if so, statement of the antecedent factors and of the consequent measures will be necessary. For example, does more identification with a parent occur when the parent's occasional absence permits the child to play the parental role? Is such identification increased to the extent of taking on symbolic self-punishment mechanisms if the socialization of the child is delayed until parental behavior has taken on strong secondarily reinforcing properties? These are questions which are also being investigated by Whiting and others with cross-cultural methods (Whiting and Child, 1953). In these ways S-R theory may assist in explicating this very important concept which the Freudians have done much to keep before us.

Sears (1943) has also provided a broad summary of the studies which can be viewed as bearing on certain psychoanalytic formulations. The objective data he presents provide a clear enumeration of the reasons why S-R theorists feel there is a need for radical reformulation of Freudian theory. Nevertheless, there is clear evidence that learning theorists have obtained much stimulation from Freudian thinking.

Attitude change. S-R theorists have long been interested in the measurement and function of attitudes, but Hovland, *et al* (1949) have restimulated interest in the study of attitude *change,* the area where learning theory is probably most directly relevant at present. Whereas most studies have shown that the effect of some educative effort directed toward verbal attitudes usually is greatest immediately after the attempt at change, Hovland uncovered some cases of "sleeper effect." Here, with some items, the effect became most pronounced some time *after* the educative effort.

A stimulating set of papers has followed, attempting to unravel the conditions for a sleeper effect, and to place the phenomenon within a larger theoretical framework (Hovland and Weiss, 1952; Weiss, 1952; Weiss, 1953). One hypothesis advanced was that the subjects may have felt doubtful regarding the *source* of those items which later show a sleeper effect. This doubt would lead to temporary avoidance of use of the material (nonacceptance), but the doubt would be forgotten sooner than the material itself, if the material had been well learned. The result would be a delayed effectiveness of the communication. Hovland and Weiss (1952) have confirmed some of the implications of this hypothesis, and Weiss (1953) has recently set up an experiment which *controlled* the acquisition of the content as well as explicitly controlling the discounting factor, with similar results. In short, there has been an increasing control of the variables involved and an increase in the explicit use of S-R theory as the search has gone on. These researches may give social psychology the explicit and general theory of attitude development and extinction that has long been needed to keep up with the development of measurement technique, which has been used for more limited applied purposes.

"Meaning." Osgood (1952) has recently summarized the general S-R attempts to study "meaning" with objective methods. He focuses his discussion on the problem of "mediating responses," which have loomed so importantly in Guthrie's thinking and in the theorizing of Miller and Dollard and many others. He has also developed an attempt at an independent measure of the complex matrix of responses which "meaning" implies. Hunt (1952) has recently suggested that these developments may change drastically our methods of measuring public opinion and also our techniques of clinical diagnosis. Further, since the measures Osgood is developing are placed in the context of general S-R theory, the implications of the measures are broad and are quickly testable. This departure represents another attempt to obtain a valid measure of a complicated response-side structure and then to relate it with behavior and, presumably, to antecedent conditions. It may well be one of the most important of the large number of recent endeavors in systematic generalization of S-R theory.

Other S-R theorists are interested in this same area but are more concerned with the development of the methods of "information" analysis which have been summarized by George Miller (1951). In this article Miller also discusses aspects of the unpublished William James Lectures on language behavior given by B. F. Skinner at Harvard in 1947, and touched upon in his recent book (Skinner, 1953).

Other generalizations. This volume will attest in other contexts the degree to which S-R concepts have been generalized in terms of limited hypotheses or as larger systems. Let us briefly list an additional sample of these attempts. Mitrano's (1939) extension of Hullian concepts to the behavior of children stands as an exploratory attempt at systematic operational extension to which the work of Miller and Dollard (1942) and of Solomon and his co-workers should be added (Lambert, Solomon, and Watson, 1948; Walk, 1952; Minturn, 1952; and Olds, 1952). Solomon's students have also worked on operational extensions of the rich controversies of the animal laboratory to the behavior of human adults, and with Howes (1951), Solomon has developed a language response-probability theory for handling certain data in the area of the "new look" in perception, relating speed of recognition to certain "value" factors. Mowrer's version of a dual theory has motivated Solomon and Wynne (1950) to an attempt to vary the condition of the autonomic nervous system in dogs by use of drugs and surgery and to study the effect of these changes on avoidance learning. Such experimentation has led to the important suggestion that certain kinds of anxiety learning may be unidirectional, i.e., a fear of the dentist based on strong pain conditioning may never be completely extinguishable. Such a principle, if further upheld, would lead to important practical implications in setting the limits on some kinds of therapy. Some practical suggestions for therapy have been interlarded in a recent report of these studies with dogs (Solomon, Kamin, and Wynne, 1953).

The outline by F. H. Allport (1942) of a descriptive language based upon S-R thinking for handling the problem of "collective behavior" cannot be adequately discussed here. He has outlined the properties of an "event system" based on the individualist position which he has brilliantly defended for many years (F. H. Allport, 1924, 1942). The research possibilities of this model are now being worked out (Vallance, 1951). The conception is probably the most inclusive theory for discussing broad problems of collective behavior yet advanced by a psychologist and may provide an explicit denotative referent for the notion of collective behavior, in a sense which most sociologists and anthropologists have failed to do. As far as the notion of "total social system" is concerned, Allport suggests that such an assumption be dropped and the degree of systematization be left for empirical discovery by use of his descriptive techniques.

McClelland (1951) has presented the results of a long series of ingenious studies in an attempt to develop an objective measure of achievement needs in the members of several cultures. Work on the antecedent factors involved in the maximization of this learned drive has hardly begun, but work on the reliability and validity of the response-side measures is well advanced. Others are working on adequate measures for affiliative drives (Shipley and Veroff, 1952), and work on both S-R and R-R analysis of aggressive drives has continued as a lasting interest of social psychologists since the publication of the frustration-aggression hypothesis by the Yale group (Dollard *et al.*, 1939).

J. J. Gibson (1952) has recently reanalyzed the autokinetic experiments in terms of a learning theory developed for "perceptual learning" which has clear S-R implications. Gibson suggests that a focus on the possibilities of specifying the conditions under which human beings come to respond differentially to more and more abstract and complex dimensions of stimuli may lead to less emphasis on postulating intervening variables. Such a suggestion may be readily accepted by some S-R psychologists who have long held that the current proliferating lists of intervening variables are rarely worth the trouble, and that some current uses of the notion of stimulus are unnecessarily vague.

RELATION TO OTHER BEHAVIOR SCIENCES

Learning theory has intrigued social sicentists who are dealing with problems of the fields of sociology, anthropology, education and psychiatry. It may be recalled from the diagram of Hull's system that there is explicit reference to environmental factors on both the antecedent and the intervening area. Complication of the diagram can be used to represent the "situation," social or otherwise, and the intervening matrix can be complicated to describe the relatively lasting effects of past experience. The basic machinery is thus provided for translation into other languages of the social sciences. Culture, for example, has often been defined in terms of habituation, and "habit strength" can therefore be expanded to explore the notion of "culture strength," and the notion of the "individual" in the diagram can be expanded to represent a "modal individual" at some stated level of abstraction. Any particular interpretation into S-R theory is, of course, not necessarily acceptable to the original user of the concept.

It should also be recalled that S-R theorists have long focused on the relatively limited problem of stating the conditions for *change* of structures. Since they are, therefore, not usually ego-involved, at least by theoretical preconception, in any *particular* categorization of structures or "habit family hierarchies," this opens up the system for trying out those notions of structure developed by other social sciences. The notion of "role" as a categorizing or unit concept has long been in use by S-R theorists, as have "culture," "interaction," "meaning of the situation," and many other concepts of social science.

Several anthropologists have found S-R theory useful. Murdock (1949) in his cross-cultural analysis of taboos has found the concept of stimulus generalization of value in his explanations. Kluckhohn in collaboration with Mowrer (1944) has provided a broad outline for the study of personality development in the context of culture. Gillen (1948) and Dollard (see Miller and Dollard, 1942) have both found it possible to integrate S-R concepts into the conception of anthropological problems. Whiting's *Becoming a Kwoma* (1941) makes considerable use of S-R theory and also suggests some alternative lines of development of the theoretical system itself. Whiting, in collaboration with Child (1953), has, through the integration of S-R theory with certain Freudian concepts, laid the foundation for a broad use of anthropological materials in the testing of hypotheses regarding the effects of socialization practices on personality development.

Davis and Havighurst (1946) have worked with S-R reinforcement notions in an analysis of the differences between certain social classes in the United States. Their methods and approach have added to the generality of some learning principles and possibly they have enlarged the concept of class by giving it a denotation in terms of specific kinds of interaction of child with parent which provide conditions for differential kinds of learning to occur. The Mead tradition in social psychology, as represented by such theorists as Cottrell (1942) and Foote (1951), has developed independently of most of the S-R theories which we have emphasized in this chapter. Possibilities for a reintegration of these disparate streams of behavior theory are suggested in the role theorists' reawakened interest in problems of motivation (Foote, 1951) and in certain clear parallels which are discernible in the conceptualizing of Cottrell (1942) on the one hand and of Sears (1951) and Miller and Dollard (1941) on the other.

Sears (1951) has suggested, as a result of his active attempts to extend S-R theory, that the *dyadic framework* be taken for the analysis of all social interaction and personality development problems. He refers to the fact that most learning theory has been stated in terms of the behavior of a single individual or of a "modal" person and that the dyadic or two-person interaction concept has been relatively de-emphasized. It is possible, however, as he points out, to derive from S-R theory (as well as other S-R theories) such dyadic interaction since, as we have pointed out, S-R theory leaves considerable place for the effects of the social as well as the physical environment. Sears feels that the notion of the dyad can be made more clear by the use of a new intervening variable which he symbolizes as $r_{ee}s_{ee}$.

The diagram in Fig. 2 should make this concept more clear. The symbol "$r_{ee}s_{ee}$" refers to the fact that not only does the *goal response*

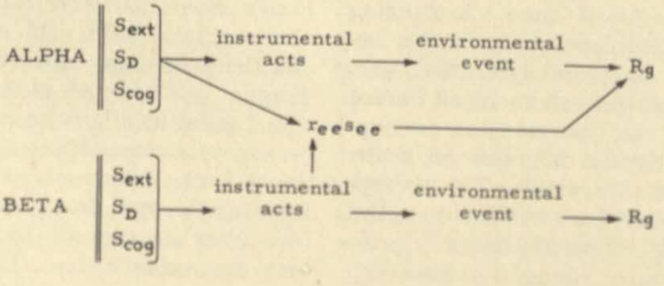


FIG. 2. The dyadic sequence with anticipatory responses to the environmental event ($r_{ee}s_{ee}$). (From Sears, 1951, p. 471)

fractionate and come to anticipate the occurrence of the goal, but that those *instrumental acts which have preceded the environmental event* leading to the possibility of the goal response also fractionate and come to anticipate the overt instrumental acts themselves. Therefore the reference is to the fact that in social interaction there is usually involved the expectation on the part of alpha of certain behaviors on the part of beta which provide alpha with the possibilities of an "Rg," or drive reduction. This concept appears to be quite close to Mead's notion of the "expectation of the other" (Mead, 1934), and there is no reason why modern S-R reinforcement theory cannot consider such symbolic expectations as one of the important determinants of social behavior. Further explication of Sears' new concept may lean upon hunches from the sociological side of social psychology, which have been developed, for example, by Lindesmith and Straus (1949), as well as by Cottrell. Although conceptual intertranslation is not a necessary prerequisite for scientific advance, it can lead to expansion of the energy and thought thrown upon common problems. It is heartening that such independent conceptual concurrence can occur in the fledgling field of social psychology.

CONCLUSION

We have sketched in broad, but probably selective, fashion some of the current endeavors to sound out the implications of the general stimulus-response position and have outlined some of the attempts to construct increasingly explicit hypotheses regarding human interaction. It is peculiarly difficult to encapsulate the S-R tradition in easy phrases, as of 1953. Change, broadening, greater precision, and extension are all the order of the day, and there is perhaps as much stimulating disagreement *among* S-R theorists as there is between S-R theorists and other orientations.

Whatever the future of this highly cognitive, deeply conative but nonmentalistic psychology may be, there is little doubt that there is still much to be heard from its adherents. A tradition is in good hands as long as there remains an active searching for defining data, experimental control, precise thought, and, perhaps above all, vulnerability of formulation.

References

Allport, F. H. *Social psychology.* Cambridge: Houghton Mifflin, 1924.

Allport, F. H. An event-system theory of collective action: with illustrations from economic and political phenomena and the production of war. *J. soc. Psychol.,* 1940, *11,* 417-445.

Allport, F. H. Methods in the study of collective action phenomena. *J. soc. Psychol.,* 1942, *15,* 165-185.

Allport, G. W., & Vernon, P. E. *Studies in expressive movement.* New York: Macmillan, 1933.

Bergmann, G. Psychoanalysis and experimental psychology: a review from the standpoint of scientific empiricism. *Mind,* 1943, *52,* 122-140.

Bergmann, G., & Spence, K. W. The logic of psychophysical measurement. *Psychol. Rev.,* 1944, *51,* 1-24.

Boguslavsky, G. W. Interruption and learning. *Psychol. Rev.,* 1951, *58,* 248-255.

Boguslavsky, G. W. A mathematical model for conditioning. Paper read at E.P.A., Boston, 1953.

Brogden, W. J. Animal studies of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 568-612.

Bronfenbrenner, U. Personality. In C. P. Stone (Ed.), *Annual review of psychology.* Vol. 4. Stanford: Annual Reviews, Inc., 1953. Pp. 157-182.

Brown, J. S., & Farber, I. E. Emotions conceptualized as intervening variables — with suggestions toward a theory of frustration. *Psychol. Bull.,* 1951, *48,* 465-495.

Bruner, J. S., & Goodman, C. C. Value and need as organizing factors in perception. *J. abnorm. soc. Psychol.,* 1947, *42,* 37-39.

Bruner, J. S., & Postman, L. Emotional selectivity in perception and reaction. *J. Pers.,* 1947, *56,* 69-77.

Bruner, J. S., & Postman, L. Perception, cognition and behavior. *J. Pers.,* 1949, *18,* 14-31.

Brush, F. R., Bush, R. R., Jenkins, W. O., John, W. F., & Whiting, J. W. M. Stimulus generalizations after extinction and punishment: an experimental study of displacement. *J. abnorm. soc. Psychol.,* 1952, *47,* 633-640.

Bush, R. R., & Mosteller, F. A model for stimulus generalization and discrimination. *Psychol. Rev.,* 1951, *58,* 413-423.

Bush, R. R., & Whiting, J. W. M. On the theory of psychoanalytic displacement. *J. abnorm. soc. Psychol.,* 1953, *48,* 261-272.

Child, I. L., & Waterhouse, I. K. Frustration and the quality of performance: I. A critique of the Barker, Dembo, and Lewin experiment. *Psychol. Rev.,* 1952, *59,* 351-362.

Clark, R. A. The projective measurement of experimentally induced levels of sexual motivation. *J. exper. Psychol.,* 1952, *44,* 391-399.

Cottrell, L. S. The analysis of situational fields in social psychology. *Amer. sociol. Rev.*, 1942, *7*, 370-382.

Davis, A., & Havighurst, R. J. Social class and color differences in child-rearing. *Amer. sociol. Rev.*, 1946, *11*, 698-710.

Dewey, R., & Humber, W. J. *The development of human behavior.* New York: Macmillan, 1951.

Dinsmoor, J. A. The effect of hunger on discriminated responding. *J. abnorm. soc. Psychol.*, 1952, *47*, 67-72.

Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. R. *Frustration and aggression.* New Haven: Yale Univ. Press, 1939.

Dollard, J., & Miller, N. E. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.

Doob, L. W. *Public opinion and propaganda.* New York: Holt, 1948.

Doob, L. W. *Social psychology.* New York: Holt, 1952.

Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, *57*, 94-107.

Feigl, Herbert. Operationism and scientific method. *Psychol. Rev.*, 1945, *52*, 250-259.

Foote, N. N. Identification as the basis for a theory of motivation. *Amer. sociol. Rev.*, 1951, *16*, 14-21.

French, J. R. P., Jr. Organized and unorganized groups under fear and frustration. *Authority and frustration: studies in topological and vector psychology III. Univ. Ia. Stud. Child Welfare*, 1944, *20*, 231-307.

Gibson, Eleanor J. Intra-list generalization as a factor in verbal learning. *J. exper. Psychol.*, 1942, *32*, 185-200.

Gibson, Eleanor J. The role of shock in reinforcement. *J. comp. physiol. Psychol.*, 1952, *45*, 18-30.

Gibson, J. J. Social perception and the psychology of perceptual learning. In press.

Gillen, J. *The ways of men.* New York: Appleton-Century-Crofts, 1948.

Guthrie, E. R. *The psychology of learning.* New York: Harper, 1935.

Guthrie, E. R. *Psychology of human conflict.* New York: Harper, 1937.

Guthrie, E. R. Personality in terms of associative learning. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald Press, 1944. Pp. 49-68.

Guthrie, E. R. *The psychology of learning.* Rev. Ed., New York: Harper, 1952.

Guthrie, E. R., & Horton, C. P. *Cats in a puzzle box.* New York: Rinehart, 1946.

Hebb, D. O. The role of neurological ideas in psychology. *J. Pers.*, 1951, *20*, 39-55.

Hilgard, E. R. *Theories of learning.* New York: Appleton-Century-Crofts, 1948.

Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 517-567.

Hollenberg, Eleanor, & Sperry, Margaret. Some antecedents of aggression and effects of frustration in doll play. *J. Pers.*, 1951, *1*, 32-43.

Homans, G. C. *The human group.* New York: Harcourt, Brace, 1950.

Hovland, C. I. Human learning and retention. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 613-689.

Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. *Experiments on mass communication.* Princeton: Princeton Univ. Press, 1949.

Hovland, C. I., & Weiss, W. The influence of source credibility on communication effectiveness. *Publ. Opin. Quart.*, 1952, *15*, 635-650.

Hull, C. L. *Principles of behavior.* New York: D. Appleton Century, 1943.

Hull, C. L. Behavior postulates and corollaries — 1949. *Psychol. Rev.*, 1950, *57*, 173-180.

Hull, C. L., Hovland, C. I., Ross, R. T., Hall, M., Perkins, D. T., & Fitch, F. B. *Mathematico-deductive theory of rote learning.* New Haven: Yale Univ. Press, 1940.

Hunt, J. McV. Psychological services in the tactics of psychological science. *Amer. Psychologist*, 1952, *7*, 608-622.

Jenkins, W. O., & Sheffield, F. D. Rehearsal and guessing habits as sources of the spread of effect. *J. exper. Psychol.*, 1946, *36*, 316-330.

Jenkins, W. O., & Stanley, J. C., Jr. Partial reinforcement: a review and critique. *Psychol. Rev.*, 1950, *47*, 193-234.

Kimble, G. A. Conditioning as a function of the time between conditioned and unconditioned stimuli. *J. exper. Psychol.*, 1947, *37*, 1-15.

Klein, G. S., & Krech, D. The problem of personality and its theory. *J. Pers.*, 1951, *20*, 2-23.

Koch, Sigmund. The logical character of the motivation concept. I., *Psychol. Rev.*, 1941, *48*, 15-38.

Lambert, W. W. Acquisition and extinction of instrumental response sequences in children and in rats. Unpublished doctoral dissertation, Harvard University, 1950.

Lambert, W. W., Solomon, R. L., & Watson, P. O. Reinforcement and extinction as factors in size estimation. *J. exper. Psychol.*, 1949, *39*, 637-641.

Lawrence, D. H. The transfer of a discrimination along a continuum, *J. comp. physiol. Psychol.*, 1952, *45*, 511-516.

Leeper, R. W. A motivational theory of emotion to replace "emotion as disorganized response." *Psychol. Rev.*, 1948, *55*, 5-21.

Leeper, R. W. Book review — *Essentials of behavior* by Clark L. Hull. *Amer. J. Psychol.*, 1952, *65*, 478-491.

Lewin, K. *A dynamic theory of personality.* New York: McGraw-Hill, 1935.

Lindesmith, A. R., & Strauss, A. L. *Social psychology.* New York: Dryden, 1949.

McClelland, D. C. *Personality.* New York: William Sloane, 1951.

MacLeod, R. B. The phenomenological approach to social psychology. *Psychol. Rev.,* 1947, *54,* 193-210.

McReynolds, Paul. Logical relationships between memorial and transient functions. *Psychol. Rev.,* 1950, *57,* 140-144.

Maier, N. R. F. *Frustration: the study of behavior without a goal.* New York: McGraw-Hill, 1949.

Marx, M. H. *Psychological theory. Contemporary readings.* New York: Macmillan, 1951.

Mead, G. H. *Mind, self and society.* Chicago: Univ. Chicago Press, 1934.

Meehl, P. E. On the circularity of the law of effect. *Psychol. Bull.,* 1950, *47,* 52-75.

Miller, G. A. *Language and communication.* New York: McGraw-Hill, 1951.

Miller, N. E. Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald, 1944. Pp. 431-465.

Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.,* 1948, *43,* 155-178.

Miller, N. E. Comments on theoretical models illustrated by the development of a theory of conflict. *J. Pers.,* 1951a, *20,* 82-100.

Miller, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951b. Pp. 435-472.

Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.

Minturn, L. Latent learning and discrimination reversal in human discrimination learning. Paper read at Eastern Psychol. Assn., Boston, April, 1953.

Mitrano, A. J. Principles of conditioning in human goal behavior. *Psychol. Monogr.,* 1939, *51,* No. 4.

Moore, O. L., & Lewis, D. J. Learning theory and culture. *Psychol. Rev.,* 1952, *59,* 380-388.

Mosteller, F., & Bush, R. R. Paper read at Brown University Colloquium, January, 1950.

Mosteller, F., & Bush, R. R. A mathematical model for simple learning. *Psychol. Rev.,* 1951, *58,* 313-323.

Mowrer, O. H. A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.,* 1939, *46,* 553-565.

Mowrer, O. H. On the dual nature of learning — a reinterpretation of "conditioning" and "problem-solving." *Harvard educ. Rev.,* 1947, *17,* 102-148.

Mowrer, O. H. *Learning theory and personality dynamics.* New York: Ronald Press, 1950.

Mowrer, O. H. Two-factor learning theory: summary and comment. *Psychol. Rev.,* 1951, *58,* 350-354.

Mowrer, O. H. Secondary reinforcement in two-factor learning theory. Unpublished manuscript, 1952.

Mowrer, O. H., & Kluckhohn, C. Dynamic theory of personality. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* Vol. 1. New York: Ronald Press, 1944. Pp. 69-135.

Murdock, G. P. *Social structure.* New York: Macmillan, 1949.

Murray, H. A. *Explorations in personality.* New York: Oxford Univ. Press, 1938.

Nowlis, V., & Nowlis, Helen H. Child Psychology. In C. P. Stone (Ed.), *Annual review of psychology.* Vol. 3. Stanford: Annual Reviews, Inc., 1952. Pp. 1-28.

Olds, J. The acquisition of motives. Unpublished doctor's dissertation, Harvard Univ., 1952.

Osgood, C. E. The nature and measurement of meaning. *Psychol. Bull.,* 1952, *49,* 197-237.

Parsons, T., & Shils, E. A. Values, motives and systems of action. In T. Parsons & E. A. Shils (Eds.), *Toward a general theory of action.* Cambridge: Harvard University Press, 1951. Pp. 47-275.

Pavlov, I. P. *Conditioned reflexes.* (Transl. by G. V. Anrep). London: Oxford Univ. Press, 1927.

Postman, L. The history and present status of the law of effect. *Psychol. Bull.,* 1947, *44,* 489-563.

Queener, E. L. *Introduction to social psychology.* New York: Sloane, 1951.

Schlosberg, H. The relationship betwen success and the laws of conditioning. *Psychol. Rev.,* 1937, *44,* 379-394.

Schlosberg, H., & Solomon, R. L. Latency of response in a choice discrimination. *J. exper. Psychol.,* 1943, *33,* 22-39.

Sears, Pauline S. Level of aspiration, academically successful and unsuccessful children. *J. abnorm. soc. Psychol.,* 1940, *35,* 498-536.

Sears, R. R. *Survey of objective studies of psychoanalytic concepts.* New York: Social Science Research Council, 1943.

Sears, R. R. Experimental analysis of psychoanalytic phenomena. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* Vol. 1. New York: Ronald, 1944. Pp. 306-332.

Sears, R. R. Ordinal position in the family as a psychological variable. *Amer. sociol. Rev.,* 1950, *15,* 397-401.

Sears, R. R. Social behavior and personality development. In T. Parsons & E. A. Shils (Eds.), *Toward a general theory of action.* Cambridge: Harvard Univ. Press, 1951. Pp. 465-478.

Seward, J. P., & Levy, N. Sign learning as a factor in extinction. *J. exper. Psychol.,* 1949, *39,* 660-669.

Sheffield, F. D., & Roby, T. B. Reward value of a non-nutritive sweet taste. *J. comp. physiol. Psychol.,* 1950, *40,* 287-298.

Sheffield, F. D., Wulff, J. J., & Backer, R. Reward value of copulation without sex drive reduction. *J. comp. physiol. Psychol.,* 1951, *44,* 3-8.

Shipley, T. E., & Veroff, J. A projective measure of need for affiliation. *J. exper. Psychol.,* 1952, *43,* 349-356.

SKINNER, B. F. *The behavior of organisms.* New York: Appleton-Century-Crofts, 1938.

SKINNER, B. F. The operational analysis of psychological terms. *Psychol. Rev.,* 1945, *52,* 270-277.

SKINNER, B. F. Are theories of learning necessary? *Psychol. Rev.,* 1950, *57,* 193-216.

SKINNER, B. F. *Science and human behavior.* New York: Macmillan, 1953.

SMITH, S. *The psychological origin and treatment of enuresis.* Seattle: Univ. Washington Press, 1948.

SMITH, W. I., & ROSS, S. The social behavior of vertebrates: a review of the literature (1939-1950). *Psychol. Bull.,* 1952, *49,* 598-627.

SOLOMON, R. L. The influence of work on behavior. *Psychol. Bull.,* 1948, *45,* 1-40.

SOLOMON, R. L., & HOWES, D. H. Word frequency, personal values, and visual duration thresholds. *Psychol. Rev.,* 1951, *58,* 256-270.

SOLOMON, R. L., KAMIN, L. J., & WYNNE, L. C. Traumatic avoidance learning: the outcomes of several extinction procedures with dogs. *J. abnorm. soc. Psychol.,* 1953, *48,* 291-302.

SOLOMON, R. L., & WYNNE, L. C. Avoidance conditioning in normal dogs and in dogs deprived of normal autonomic functioning. *Amer. Psychologist,* 1950, *5,* 264.

SPENCE, K. W. The methods and postulates of "behaviorism." *Psychol. Rev.,* 1948, *55,* 67-78.

SPENCE, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 690-729.

SPENCE, K. W. Mathematical formulations of learning phenomena. *Psychol. Rev.,* 1952a, *59,* 152-160.

SPENCE, K. W. Clark Leonard Hull: 1884-1952. *Amer. J. Psychol.* 1952b, *65,* 639-646.

STEVENS, S. S. (Ed.) *Handbook of experimental psychology.* New York: Wiley, 1951.

STROUD, J. B. Psychology in education. New York: Longmans, Green, 1946.

THORNDIKE, E. L., *et al.* An experimental study of rewards. *Teach. Coll. Contr. Educ.,* 1933, No. 580.

TINBERGEN, N. *The study of instinct.* Oxford: Clarendon Press, 1951.

TOLMAN, E. C. *Purposive behavior in animals and men.* New York: Appleton-Century-Crofts, 1932.

TOLMAN, E. C. A psychological model. In T. Parsons & E. A. Shils (Eds.), *Toward a general theory of action.* Cambridge: Harvard Univ. Press, 1951. Pp. 279-364.

VALLANCE, T. R. Methodology in propaganda research. *Psychol. Bull.,* 1951, *38,* 32-61.

WALK, R. D. Effect of discrimination reversal on human discrimination learning. *J. exper. Psychol.,* 1952, *44,* 410-419.

WATSON, J. B. Psychology as the behaviorist views it. *Psychol. Rev.,* 1913, *20,* 158-177.

WEISS, W. A study of a "sleeper" effect in opinion change. Unpublished doctor's dissertation, Yale University, 1952.

WEISS, W. A "sleeper" effect in opinion change. *J. abnorm. soc. Psychol.,* 1953, *48,* 173-180.

WHITE, C. T., & SCHLOSBERG, H. Degree of conditioning of the GSR as a function of the period of delay. *J. exper. Psychol.,* 1952, *43,* 357-362.

WHITING, J. W. M. *Becoming a Kwoma.* New Haven: Yale Univ. Press, 1941.

WHITING, J. W. M. A brief theory of purposive acquired drives. Unpublished manuscript, 1948.

WHITING, J. W. M., & CHILD, I. L. *Child training and personality.* New Haven: Yale University Press, 1953.

WRIGHT, G. C. Projection and displacement: a cross-cultural study of the expression of aggression in folktales. Unpublished doctoral dissertation. Harvard University, 1952.

CHAPTER 3

Cognitive Theory

MARTIN SCHEERER*
University of Kansas

Cognitive theory might be expected to deal with the problem of how man gains information and understanding of the world about him, and how he acts in and upon his environment on the basis of such cognitions. Such expectations need not carry epistemological meaning; cognitive theory could still be a legitimate branch of psychology as an empirical science whose task is to discover the processes which underlie cognitions, to determine the conditions for their occurrence, and their function in behavior. Disagreements regarding the psychological laws of cognition would then reflect natural family quarrels to be resolved by experimental test. A closer inspection of contemporary systems in psychology presents, however, a different picture. Not only does cognition play a less noteworthy role in many of their formulations, but it also fails to appear as a commonly shared problem. Influential schools seem even to question whether cognitive processes have to be posited at all in a scientific psychology. This seems to be particularly the case in modern behavior theory and in kindred stimulus-response approaches. Here cognition is neither a psychological fact nor a problem for investigation; nor does it become a postulated behavior variable. For those, however, whose views are oriented around holistic, "molar" or personalistic principles, cognition constitutes a major problem area and also an essential explanatory construct (cf. the approach to social psychology by Krech and Crutchfield, 1948).

COGNITION AS AN ISSUE

The primary issue, therefore, seems not to be *how* cognition plays a role in behavior but *whether* cognition plays a role in behavior. This formulation might lead to the objection that it creates artificial issues and straw men where only different methods, interests, and objectives prevail. Spence (1950, 1951) for one, has increasingly insisted that the difference between S-R and Cognitive Theory is chiefly one of emphasis. The first focuses on observable or hypothetical S-R connections, the second on reported or inferred perceptual organizations as mediators between stimulus patterns and behavior. Thus, Spence compares S-R theory with Perception or S-S theory, which emphasizes Sign-Significate concepts. It could also be argued that some of the recent studies on communication, though chiefly emphasizing S-R aspects, do measure cognition (G. Miller, 1953). Information is here determined by the number of possible response outcomes, e.g., telling a child to find a cookie by lifting one of sixteen boxes delimits the response to sixteen possible outcomes; telling him the cookie is under one of four boxes on the left reduces the sixteen to four alternatives. Information is defined "probabilistically" by the number of reduced alternatives necessary to control the number of possible outcomes. Important as these studies are for measuring information in terms of message (S) and behavior (R), they do not actually deal with cognition as a psychological problem. Thus, information "is not synonymous with 'meaning' . . . the amount does not specify the content, value, truthfulness, exclusiveness, history, or purpose of the information. . . . Nothing is said about whether the information is understood or believed" (G. Miller, 1953, pp. 3, 4).

Even if such differences in approach were taken to be merely matters of emphasis, the question still remains why one approach so studiously avoids semantic reference to cognitive factors while another seems to consider it indispensable. Is it because cognition hypotheses are often couched in apparently introspective terms?

The influence of operationism and behaviorism. The rising prominence of neopositivism might be thought responsible for the above "differences in emphasis" between S-R and Cognitive Theory. In physics, Bridgman's recognition that concepts such as length or distance do not mean the same when used in different systems led to the method of "operationism" which stipulates that the observational or experimental operation should be specified in the definition of every concept. In this sense concepts do not denote absolute or existential properties but are constructs of the observer. Stevens (1939) has most eloquently championed this method for the discipline of psychology; its propositions can have validity only if anchored in concrete operations which can be performed, are repeatable and public. This rule might have fostered a tendency to interpret reference to cognition as reference to "private" data of experience. Thus it has been charged that such terms as insight, understanding, conceptual thinking, phenomenal field, and self-image are "mentalistic," i.e., contaminated with the subjectivity of consciousness.

---

*My sincere thanks go to Dr. Beatrice Wright, Aaron Hershkowitz, and Dr. Lee Meyerson for their patient and constructive criticism of the wording of this chapter.

Yet operationists have not concurred in a common verdict against "phenomenological" description. Its use in terms of verbal report is legitimate according to Boring (1945) and others, e.g., when subjects discriminate between stimuli. Stevens holds that even " 'mentalistic' concepts like percept, image, and idea can be operationally defined." And Spence (1944), though skeptical of the predictive value of the phenomenological method, accepts its use of verbal reports about perceptual contents, for instance in the experiments on figural aftereffects by Köhler and Wallach. These kinds of "perceptual responses" can lead to "response-inferred" theoretical constructs such as life-space. On the other hand, there is still an unsettled debate over the question of how to deal in operationism with the basically private mode of direct experience in its unverbal, preverbal, or verbal aspects (Bridgman, 1945; Skinner, 1945; Pratt, 1945a, b). This uncertainty might account for the constraint in the research interests of operational psychologists who have rarely concerned themselves with clinical and social problems. But it does not explain their self-imposed restrictions in the use of cognitive constructs within other problem areas which are less person-centered as, for example, sense perception or animal and human learning. The notable exception of postulated "cognitive maps in rats and men" in Tolman's "operational behaviorism" demonstrates that such intervening variables in learning need not be objected to on operational grounds.

Operationism as a *method*, therefore, was not the exclusive source of resistance to a more universal inclusion of cognitive factors in the making of psychological theory. Rather, this resistance among psychologists dates back to the impact of Watson's Behaviorism. In the "Zeitgeist" of Positivism and Reflexology, psychological data became more and more limited to the observable facts of overt behavior. Only when questions of consciousness and thinking could be translated into motor terms was their status as legitimate scientific problems recognized. Paradoxically, as Hilgard (1948) pointed out, it was those students of the doctrine of "association of ideas" who became the advocates of equating ideas and consciousness with motor activity. These tenets prepared the ground for the ready acceptance of Neo-Positivism, which tended to anchor operational methods in physics and mathematics as the foundation for a "unity of science." The influence of Neo- or Logical-Positivism is reflected in the modern behavior theories of Hull and Skinner. Both formulate physical stimulus-response relationships either in terms of mathematical or empirical functions. Basically both attempt to make their concepts synonymous with concrete operations, in conformance with the new concept of the "unity of science."

Boring (1945) aptly summed up this entire historical development and the common philosophic directive of Neo-Behaviorism and Neo-Positivism in the following statement: "All positivism from Mach to Bridgman has stressed the advantage for clarity of reducing dubious concepts to empirical ultimates — to the data of observation (Mach) or the processes of observation (Bridgman)," (p. 244). For most S-R theorists, empirical ultimates are defined as behavior data linked to associative neural connections. Thus, empirical ultimates came to mean physical response units, and the processes of observation, physical measurements. This identification made cognition, because it eludes physical concepts, a rather "intangible" problem. Cognition theorists are no less empirically minded in the search for empirical ultimates. But for them these ultimates include phenomenal data as long as they are tied to objective procedures. Whether or not such data of direct experience correspond to the underlying processes, and whether or not they are ultimately reducible to accompaniments of nerve impulses, they are held to be a primary source of information. It is granted that the examination of direct experience cannot be the sole method of psychological investigation. Phenomenological description is nevertheless considered a necessary step in the study of such problems as the physiological correlates of emotional *contents*. This step seems also necessary for the discovery of "rationalization," where the discrepancy between introspective report and actual behavior is crucial. Asch's comment on the functional relationship between experiencing and acting typifies the entire position: "To limit investigation to the observation of action alone would be to ignore the paramount fact that the actor is constantly registering in awareness what is happening to him and that this alters his subsequent acts" (1952, p. 68).

PREMISES DIFFERENTIATING S-R AND COGNITIVE THEORY

A further methodological difference between representative S-R and cognitive theorists concerns the selection of empirical ultimates as to their complexity — whether physical or nonphysical. The S-R advocates reason that it would be "premature" to study the more complex from the start; one can deal later with complex phenomena by treating them as a combination of the simpler. S-R connections are selected for study because they are considered the simplest behavior units. Most cognition advocates, on the other hand, question the adequacy of beginning with simple S-R connections. They reason that it would be "premature" to base explanatory concepts on such connections because these have been isolated from the total complex behavior. It is asserted that behavior is characterized by organization, and it is doubted that a theory of total behavior patterns can be based upon a mere combination of previously isolated simpler ones. A qualitative analysis of the perceptual aspects of the stimulus and of the response in their *total* setting should precede quantification. It is further asserted that cognition is a part of behavior organization; therefore, it also plays a role in determining the S-R units which cannot be dealt with by themselves.

The holistic or molar aspect of organization has been made a critical focus of theory and method by such men as Stern (1938), G. W. Allport (1937), Goldstein (1939), Angyal (1941), Tolman (1932), MacLeod (1947), Sloane (1945), Leeper (1951), and, of course, the founders of the Gestalt School. They have pointed out several relevant preconceptions which have contributed to the development of S-R theories without the postulation of cognitive factors. Not all of these preconceptions have been made explicit. They become apparent when S-R constructs (reflex, drive, habit, etc.) are scrutinized. Each of the aforementioned authors has analyzed one or another of these premises. Their criticisms of these preconceptions can be summarized in the following statements, without regard for their historical origin or order.

Reductionism: the "attempt to explain a complex interrelated whole in terms of its simpler elements or parts, or in terms of elements belonging to a lower level of phenomena" (Sloane, 1945, p. 217). (This takes on different forms, which Sloane considers to be the fallacies of isolation, constants, mathematical summation, origin, and of metaphor or analogy.)

The atomistic reductive bias: the belief that something small is more fundamental than something large (MacLeod, 1947).

The genetic bias: the belief that that which is genetically early is more fundamental than that which occurs later (MacLeod, 1947).

Peripheralism: "a viewpoint or faith that holds that psychological processes fundamentally may be conceived as relatively simple or mechanical links between peripheral events or receptor stimulations at one end and effector activities at the other end. It recognizes the existence of nervous processes . . . but it is convinced that the apparent complexity of the nervous processes must ultimately be reducible to relatively simple connections" (Leeper, 1946, p. 3).

The reinforcement bias: the assumption of specific primary drives or derived needs which activate behavior; that no learning occurs without the reduction of these drives or needs (G. W. Allport, 1937; Goldstein, 1939).

Clearly, these preconceptions overlap and are interrelated. When taken together they reflect an elementarism which, according to cognitive theorists, is incompatible with the concept of organization that includes cognition as a central factor. As Allport expressed it, "Addiction to machines, rats, or infants leads us to overplay those features of human behavior that are peripheral, signal-oriented, or genetic. Correspondingly, it causes us to underplay those features that are central, future-oriented, and symbolic" (1947, p. 190). In addition, Allport speaks here of a "machine model" that is being followed by psychologists. This may be exemplified by the "robot" paradigm of Boring (1946) to which he tried to assimilate human behavior, and by Hull's (1943, p. 384) definition of the organism as a "more or less successfully self-maintaining mechanism" — mechanism being defined as "physical aggregate."

The foregoing critical formulations of S-R premises have, in turn, been criticized as expressing bias by attaching emotional values to certain methods rather than attaching value to scientific objectivity (Marx, 1951; Spence, 1948). What is more crucial, however, is that these formulations bear on major issues involving cognitive theory — issues that have arisen in actual research. In what follows these issues will be discussed in the context of three problem areas: molar-molecular, conation and cognition, cognitive development and genetic reductionism.

MOLAR-MOLECULAR

Tolman (1932) introduced the terms *molar* and *molecular* to contrast the descriptive units of his behaviorism with those of Watson. Watson's approach is called molecular because it defines behavior in terms of its strict underlying physical and physiological details (i.e., receptor-conductor-effector processes *per se*). For Tolman, behavior is an emergent phenomenon that has descriptive properties of its own, namely, molarity qua behavior-acts as whole-responses. All molar properties of behavior acts must have physiological correlates that will have to be discovered. But it is claimed that at the present state of our knowledge behavior acts have qualities different from the physiological facts which make them up, and therefore cannot be deduced directly from such molecular elements. At the same time, Tolman pointed out that Watson himself did not remain consistent in his molecular definition of stimulus and response. On the one hand, Watson defined the stimulus in terms of light waves, sound waves, or gaseous particles, etc., and the response in terms of smooth or striped muscle twitches. On the other hand, he also had to introduce such concepts as "situation" on the stimulus side and "acts" on the response side, e.g., talking, writing a letter, or bricklaying. Since these concepts are really molar, Watson "dallied with two different notions of behavior, though he himself has not clearly seen how different they are" (Tolman, 1932, p. 6). For Watson, situations and acts were, of course, reducible to a complex group of stimuli and responses. But, as Tolman argues, the respective integrations present properties not reducible to a mere enumeration of the elements, e.g., of movements qua movements of an "act."

Elaborating the molar conception in terms of observable acts, Tolman developed "objective" criteria of purposive behavior which may be summed up in the following propositions: (1) Behavior is a type of commerce with the environment that can take place only in a whole organism and not in specific sensory and motor segments which are insulated each from the other. Molar behavior thus requires mutual interconnections between all parts of the organism. (2) The complete identification of a single behavior act requires an account of "the getting to or from a specific goal-object." (3) This involves the selection of appropriate means-objects or shortest routes relative to the direction and distance of the goal-object. The last two propositions lead to the concept of docility which corresponds to the observed fact that while the organism persists through trial and error in reaching the goal, it benefits from such trials in that it can select and retain the more efficient means to that end.

On this basis cognition and purpose become "objectively definable" behavior variables. Cognition is defined as the initial orientation towards a goal-object in terms of distance, direction, and valence character, and as the selection of a means-object or path in terms of its suitability for reaching the goal. Cognition is further definable as the behavioral change upon attaining the goal and as the variation of behavior if the goal does not prove to have the initially perceived valence. Correspondingly, purpose becomes objectively definable by the fact of continuously differing reactions with an *invariant terminal* direction as, for example, in detour behavior. Inasmuch as the foregoing propositions are interrelated, purpose and cognition are considered to be complementary. Subsequently, Tolman (1936) accommodated his system to operational thinking by defining cognition as an "intervening variable," a variable which mediates between the stimulus situation (the independent variable) and the resultant behavior act (the dependent variable).

Independently, contemporary authors have also employed molar principles, although they referred to molarity by using other distinctions or terms (Krech and Crutchfield, 1948; Stern, 1938; Allport, 1937; Lewin, 1935; Goldstein, 1939). Goldstein has been in the forefront in emphasizing the holistic dependency of all behavior acts, including the simplest reflex. He has, moreover, asserted more explicitly than has Tolman that the concept of organismic part-whole relationship involves two factors. The first is the hierarchical organization of functions. Cognitive processes are considered to be a more differentiated function on a higher level than most motor behavior. They are therefore assigned the role of a central determinant in behavior, particularly in the human organism. The second factor is that cognition can become independent of purpose qua *concrete* action goals, so that it can represent the environment more broadly and more abstractly, e.g., through concepts. "Molar" psychologists also differ with Tolman in his original equating of molecular events with physical correlates of behavior. They have introduced models of physiological and neural processes which are not molecular but have molar characteristics themselves. Thus Goldstein (1939) applies the concept of figure-ground to the excitation spread in the nervous system; Lashley (1942) speaks of excitation ratios and directional patterns of neural processes; the Gestaltists' theory of isomorphism postulates brain field processes of physical Gestalt character; and Krech (1950) adopts the concept of "open systems" advanced by the biologist Von Bertalanffy (1950). Krech hypothesizes unitary dynamic systems in the brain field which underlie the "phenomenal pluralism" of perception, emoting, and thinking. The common denominator for all cognitive psychologists seems, nevertheless, to be the acceptance of the tenet that behavior organization is molar and that cognition is an indispensable ingredient of this organization.

Objections have been raised that the concept *molar* has been applied too loosely, so that its criteria overlap and cover a multiplicity of meanings. In their critique Littman and Rosen discuss seven such meanings and suggest that the terms *molar* and *molecular* be abandoned, since "it does not seem possible to purge them of their confusing connotations" (1950, p. 65). According to their criticism, it is not made sufficiently clear whether molar refers to the quantitative or the qualitative nature of events, i.e., to the size or kind of units. A case in point is the difficulty of deciding where Hull's system belongs in the molar-molecular framework, a difficulty also mentioned by Hilgard (1948, p. 82ff). It is argued, moreover, that the criterion of molarity is applied both to *total* behavior and to *level* of organization. Thus, the criterion of new emergent qualities which cannot be derived from or reduced to simpler elements has been the identifying feature of total behavior; it has also been used to identify higher levels in relation to lower levels within total behavior, and so forth. Finally Littman and Rosen express concern about the tendency to make the descriptive criteria of molarity "existential" properties instead of "instrumental" criteria of empirical method. No issue can be taken with this last epistemological reservation, but it actually has no bearing on the crucial issue of the usefulness of the molar-molecular distinction.

Each of the above criticisms seems to result from failure to recognize or to formulate systematically the criteria of behavior organization. First, as Littman and Rosen themselves point out, there is no longer a cogent reason to dichotomize physiological and behavior phenomena by positing the first as molecular. We have already referred to the attempted application of molar constructs to physiological events. Secondly, the criticism that quantitative and qualitative criteria of the descriptive units are not distinguished in the use of "molar" stems from a preconceived alternative which misses the very nature of whole-part organization. Cognitive theorists do not derive molarity from the size of the descriptive unit, but derive its "size" from the organizational structure or qualitative context to which the unit belongs. In speech, for example, the first word "do" spoken in the phrase "How do you do?" has a part (molecular) function, whereas a bride's "I *do*" at her wedding ceremony has a molar function. In ordinary conversation neither the listener nor the speaker registers lip movements in the forming of words, whereas for a lip-reading deaf person, they are molar events. Furthermore, in a neurological frame of reference, the movement sequences of the speech musculature may be defined, following Lashley, as an expression of a "generalized schema" of action which predetermines "the sequence of specific acts" (1951, p. 122).

From these paradigmatic instances, it becomes clear that the molar-molecular distinction varies with the context or the organizational level selected for study. In many scientific disciplines the shifting frames of reference for these distinctions have been well accepted; viz., in one case the astronomer focuses on lawful regularities in the motion of planets without concern about their molecular structure, in another case he may focus on this structure, for example, through spectroanalysis. In perception Koffka has illustrated this dependency of units on the selection of the level of analysis. A triangle composed of dots constitutes a whole (molar) in relation to its sides, the sides in turn are wholes in relation to the dots. Size is here conceived as the relation between the enclosing and the enclosed units, i.e., as a functional whole-part relationship. In this connection Angyall (1941) has given four possible divisions of a whole such as a tree: (1) a random piecemeal cutting leading to fragments; (2) a division according to previously fixed principles extraneously imposed, e.g., dividing the tree into inch cubes; (3) a division by abstraction of distinguishable properties, e.g., color, weight, etc.; (4) a division according to structural articulation, following the structural and functional whole-part organization, e.g., branches, roots, capillaries, leaves, etc. (cf. also Goldstein, 1939).

The emphasis on proximal or distal determination. Psychologists have defended elementarism by claiming that total concrete events are never repeated exactly, but only certain features of them. Therefore even the holist has to abstract and fractionate in order to arrive at uniformities or laws (Spence, 1948). Which of the isolated features lead to concepts that are significant depends, according to the elementarists, on the pragmatic success in the discovery of laws. Spence therefore feels that one can make use either of "a proximal or distal account of the stimulus situation, and behavior may be described either in terms of movements (muscular patterns) or in terms of gross achievements," depending upon which kinds of laws one prefers to find (1948, p. 71). Heider (1939) defines the distal and proximal account of action and perception in similar terms, yet he maintains that whether proximal or distal events are made the focus of the determination is crucial for psychological theory (p. 391).

Spence's position presents two problems for consideration. The first is that not all proximal regularities are representative of the underlying organizational determinants. Thus Köhler in his critique of traditional introspective analysis has pointed out that it can destroy the functional dependency of perceptual parts, for example in the use of the reduction screen, with the result that a certain brightness or form is considered the "true" regularity of the retinal image as a proximal event. In contrast to this isolating procedure the holists attempt to determine the genuine parts of wholes versus "fictitious" elements. The second problem is whether the proximal account of the stimulus situation can be consistently carried through without contaminating this account with distal reference.

In S-R theory the tendency is to place the focus of environmental determination proximal in space and time and to avoid distal determinants. This "tendency to glorify the skin" leads to a quasiproximal rigor. One inadvertently substitutes stimulus concepts which refer to distal objects or determinants. Köhler (1947) speaks of the "experience error" where the psychologist describes stimuli in terms of the experiential distal object, though his intention is to describe proximal stimuli. Heider (1939) speaks of a disguised focus and analyzes three such typical cases in S-R theory.

The first he borrows from Tolman's critique of the conditioned response formula. The two stimuli and the two responses are not stimuli and responses in any strict physiological sense, but are environmentally defined affairs. Thus the visual, tactual, and olfactory stimuli such as food and string will differ from occasion to occasion and remain identical only because of their environmental meanings. The responses of eating and string-pulling "do not correspond to specific and invariant sets of muscle contractions but are only identifiable through successive times in terms of environmentally nameable 'manipulations'" (p. 396). This involves a distal account.

A second case is Skinner's description of the rat pressing the lever. Although the interest here is to arrive at proximal stimulus response determinants, it is admitted that "the number of distinguishable acts on the part of the rat that will give the required movement of the lever is indefinite and very large." Therefore, instead of the intended proximal events there is substituted by Skinner "a class which is sufficiently well defined by the phrase 'pressing a lever.'" This phrase denotes distal determination.

Heider's third example concerns Hull's account of the habit family hierarchy which is to "provide the basis for a purely physical theory of knowledge." The problem is: Habits can vary in situations in which there is more than one action sequence that will lead to the attainment of a particular goal or subgoal. The solution: There is a habit family hierarchy based on *past* learning of many different action sequences, for example, of different paths in different mazes. What makes the different action sequences belong to one family and thereby interchangeable is an identical element, namely, the same goal reaction as a "fractional anticipatory goal reaction." Heider asks whether these goal reactions are identical from the point of view of proximal determination. From the position of the locomoting organism, e.g., a rat in a maze, each detour within the repertoire of the family hierarchy follows proximally different S-R action sequences. Therefore, *proximally* viewed, the end place of each detour in that repertoire cannot be identical with the learned end place in the current maze. Hull does not take this into consideration, and declares that the anticipatory goal reactions are identical when the interchangeable paths terminate in a goal which is in the same *environmental* place. The identical element is thus distally determined, although the theory is based on proximal S-R sequences.

The proximal-distal issue has been dealt with in such detail because of its deeper implications, some of which Heider points out. Insofar as S-R theory attempts to explain behavior without cognitive factors, it is faced with these basic problems: (1) to establish a system of proximal determination independent of distal environmental determinants; (2) to derive from this system the observed fact that there are behavior acts which correspond to the objects, namely, the distal stimuli in the environment; (3) that this relevant, object-adequate behavior is not due to cognitive commerce with the object but only *appears* relevant and has no intrinsic relation to the properties of the distal stimuli. The last problem is usually solved by a hypothesized process of environmentally imposed "selection," as Heider calls it. From an infinite number of possible associative bonds or conditioned reflexes between any stimuli and any responses, only certain bonds will be favored in experience. This position is apparently shared by reinforcement theorists and Ames (1951) in his experiential purposivism.

A further consequence of the quasi-proximal rigor is reflected in the conceptualization of social behavior. The definition of attitude advanced by Doob (1947) and based on Miller and Dollard includes hypotheses such as these: Attitudes are similar to other learned implicit cues or drive-producing responses along a generalization gradient. An attitude has cue value in the same sense as counting the number six is a response to the cue of five, and in turn acts as a cue to elicit seven. Thus an attitude can act as a stimulus in that it evokes learned anticipatory goal responses which have the effect of tension reduction or reward. The distal references which are here implicit but not accounted for are concealed in the assumed learning processes on which the attitudes are based. The learning as well as the evocation of the attitude by cues must have a perceptual and contextual reference. Chein (1948) therefore raises this question: Is it the perceptual *response* to the proximal cue which is tension reducing or is it the *content* of the perception? The perceptual content of what leads to reward in relation to persons or objects, e.g., to approval or prestige, transcends the presumed proximal "stimulus patterns." Thus cognitive processes qua distal references are concealed behind proximal determinants.

WHAT IS A STIMULUS COGNITIVELY CONSIDERED?

From the cognitive point of view, cognition begins with perception. Koffka (1935) asks why things looks as they do, and goes on to say, "if things look as they do because they are what they are, then perception would not contain in its very makeup a cognitive problem" (p. 78). Inspection of any dotted outline figure demonstrates that there is a tendency toward unit formation, towards organized togetherness. This *patterned continuity* does not correspond to the stimuli themselves inasmuch as they are discrete points in objective space. Perceptually, however, the dots belong together as parts of the contour of a figure. Usually two adjacent spatial regions share a *continuous* physical border. Perceived contour, however, always has a "one-sided" directional quality; it does not outline the shape of the ground but of the figure; it encloses and separates the figure area against a surrounding field seen as unformed ground which extends behind the figure. These and other countless observations of perceptual unit formation have led to the conclusion that our senses do not directly mirror the physical object in the "geographic environment." Perceptual unity does not exclusively depend upon object unity. The dotted configuration or apparent motion is a classical case in point. Conversely, object unity is not always mirrored in the *same* perceptual unity. Camouflage coloring can destroy the real and create perceptually different units. It seems true, then, "that things are what they are does not explain why they look as they look" (Koffka, 1935, p. 78). The Gestaltists have been among the first to insist that "geographic" objects cannot be directly responded to as "they are," nor act directly as stimuli. As in most Gestalt theory, this thesis is expounded in the realm of vision.

> When I see a table, this table qua table does not affect my senses at all; they are affected by processes which have their origin in the sun or an artificial source of light, and which are only modified by the table before they excite the rods and cones in our retinae. Therefore, these processes, the light waves, and not the geographical objects, are the direct causes of our perceptions . . . (Koffka, 1935, p. 79)

We thus have to recognize a distinction between the physical objects, i.e., "distal stimuli," and the direct cause of our perception, the proximal stimuli which impinge on local receptors. The proximal stimuli, e.g., light rays, air particles, belong to what Heider (1927) called the "medium" which is interposed between the distal object and the sense organ. The medium possesses neither organized integration nor segregation into specific units. "The waves of light . . . do not contain as such the slightest indication of the fact that some are reflected by parts of one physical object and others by objects in its environment" (Köhler, 1947, p. 161). It is true that to a continuous distal object there corresponds an equally continuous area on the retina. Yet the elements within this retinal image are qua stimuli "functionally as independent of one another as any one of them is of an element outside the image" (Köhler, 1947, p. 162). We have, then, the paradox that our percepts must depend on proximal stimulation, since we cannot perceive the distal objects directly, yet it is these objects that determine our orientation and behavior. More concretely, when we inspect a planet through a telescope, do we perceive the light waves, i.e., the proximal stimuli, or the distant planet directly? Actually, we see neither, we do not *see* stimuli but see *because* of stimuli. What we perceive is an organized pattern occasioned by the distribution of proximal stimuli, a pattern which happens to have some correspondence to the distal physical object. The phenomenal thing, e.g., the planet, results from a process within the organism which may be called *representation of distal objects.* The sequence of events may be summarized as follows: distal source (physical object), medium, distribution of proximal stimuli, sensory organization, response to the product of organization. The product of organization is the phenomenal unit formation which represents the distal object.

This formulation leads to the question as to what determines the organizational character of a percept. It has been dealt with by the Gestaltists in two ways: first, they postulate a dynamic self-distribution of nervous excitations which is triggered off by the proximal stimuli. This culminated in Köhler's theory of isomorphism, which assumes a *formal* correspondence between brain-field patterns and phenomenal patterns; in the last analysis the latter are believed to derive from the former. In other quarters of psychology, the isomorphic hypothesis is not held essential for the acceptance of perceptual organization as a fact. Although the data on neural correlates of perception favor a molar interpretation, other "nonmosaic" alternatives to strict isomorphism have been entertained, e.g., in Lashley's theory. Second, the Gestaltists examine the proximal stimulus conditions themselves as to their possible relevance for perceptual patterning. In so doing, Köhler concludes: "Although the local stimuli are mutually independent, they exhibit formal relations such as those of proximity and similarity" (1947, p. 167). It is clearly implied that the Wertheimer factors give rise to perceptual unit formation. The organizational response is induced by the *relations* among the independent local stimuli in terms of similarity, proximity, common fate, and good continuation. On this basis the phenomenal characteristics of the perceptual field have been further explored and experimentally substantiated. The well-known demonstrations of figure-ground articulation by Rubin and others, of whole-part dependency in terms of units and subunits, of enclosing and enclosed structures, have been almost universally accepted. *The cognitive significance of these perceptual principles may be summed up in the proposition that phenomenal organization is a cohesive structured field and that the units in this field represent distal objects of the geographic environment.*

A further question concerns the behaviorally observed fact of relative *correspondence* between the representation of the distal object and the physical object itself. How does this take place? It cannot be due to a one-to-one correspondence between distal and proximal stimuli or between proximal stimuli and cognitive representation. It is well known that for each distal stimulus there exists practically an infinite number of proximal stimuli — a fact well attested to by the work of Ames (1951). The seat of a chair remains rectangular in its perceived shape, although its retinal image is rarely so. The form of a plate appears constant, although its retinal image is nearly always distorted into an oval. We know that a man thirty feet away produces a smaller retinal image than at three feet, but the perceived size remains the same. We are referring, of course, to the phenomena of size and form constancy, which contradict the "traditional" constancy hypothesis that to every local stimulus change a corresponding change in the sensation occurs. We know, moreover, that a change in distal stimuli, even if reflected in proximal stimulus variation, may not lead to a corresponding change in the percept. A white surface under low illumination may reflect less light into the eye than a black surface under strong illumination, yet the white surface may continue to look white. These well-investigated phenomena led to the conclusion that changes in proximal stimulation, regardless of whether they are accompanied by changes in the distal stimulus object, need not produce corresponding changes in the appearance of the phenomenal thing. The phenomenal thing is conservative to the extent that it preserves largely the constancy of the real physical object despite variations in proximal stimuli. It has also been recognized that the phenomenal thing can undergo alteration without variation in either the distal or proximal stimuli. Examples of this are optical illusions and ambiguous figures in which the percept can be reorganized under the influence of internal conditions, e.g., learning, acquired set, analytic attitude, sensory fatigue, and satiation effects.

In repudiating the traditional constancy hypothesis of a strict correspondence between local stimulus and sensation the theory has to postulate that something of the physical object is preserved which makes, on the whole, for its adequate representation. For one thing, the "mosaic" of retinal projection, through convergent refraction, for example, duplicates to a large degree the geometric *relations* among the various points of the surface of the distal object. This is possible because the proximal stimulus distribution, i.e., the reflected light rays, follow these formal geometric relations along the object (Köhler, 1947, p. 167). In the words of Koffka, the perceptual representation of an object "is a dynamic map of the distant stimulus, when and inasmuch as the proximal stimulus distribution possesses such geometric characteristics as will produce a psychophysical organization similar to the one of the distant stimulus object" (1935, p. 659). To explain the phenomena of size, form, and color constancy which seem to militate against this assumption, the concept of "invariance" is introduced. Invariance refers to two factors. First, that the visual response organization is oriented towards the total stimulus field so that a given part of the field will always be responded to in relation to the rest of the field. Second, that this relationship will tend to remain invariant. For example, in the case of the constancy of a white surface under low illumination, with further light reduction the white surface will still tend to appear white rather than gray, while the illumination of the surroundings will begin to appear darker. Invariances of this kind have been investigated by Wallach (1948). Boring (1952) has dealt with invariances in perception from an operational point of view. Stevens (1951) places the constancy phenomena in relationship to the problem of transformation, in which certain relations remain invariant. Cassirer (1944b), in a brilliant discussion of the pertinent psychological data, related these constancies to the theory of groups, again stressing invariant functions in the transformation. The controversy between experience advocates and the Gestaltists has not been resolved in regard to this point. The latter, however, insist that the invariance ultimately rests on the fact that no local stimulation can determine the corresponding excitation by itself, but only in connection with the totality of the stimulus field. This implies that the dynamics of visual processes — whatever their neural characteristics are — cannot be based on point-for-point local excitations. These dynamics must be based on "translocal" functions of excited receptor areas and translocal nervous processes.

A noteworthy attempt to deal with the problem of invariances has been advanced by Bateson (1951). In line with Gestalt theory, he derives the phenomenal thing from figure-ground processes of reconstructing external relations. Bateson asserts that external events are different from internal events and that a translation from the former to the latter must take place. He treats this translation with the concept of codification, which is defined as the *substitution* of one type of event for another. For an event to be codified two conditions must be fulfilled. First, codification must be systematic. "Whatever objects or events or ideas internal to the individual represent certain external objects or events, there must be a systematic relationship between the internal and the external, otherwise the information would not be useful" (p. 168). Second, codification has to preserve the relationships for which it substitutes. "While it is impossible for a man to have inside himself a tree corresponding to the external tree that he perceives, it is possible to have internal objects or events so related to each other that their relations reflect relationships between parts of the external tree" (pp. 169-170). This is a novel approach to the problem of cognition in terms of information theory: representation becomes synonymous with codification.

It follows from the foregoing analysis that there is not one answer to the question, "What is a stimulus?," but that we have to distinguish at least three. From the point of view of immediate psychological reality the stimulus is the phenomenal "thing" *as* it represents the geographic object — whether faithfully or not. From the point of view of the underlying and mediating processes the stimulus is a total, proximal field distribution *as* it elicits translocal excitation patterns in the central nervous system. From the point of view of the geographic environment, the stimulus consists of those physical units *as* they affect the organism through the medium of proximal stimulus relations. Cognition on these sensory levels may then be defined as the perceptual representation of distal objects; it takes the form of phenomenal organization *which is centrally interposed between* the source of stimulation and the behavioral adjustment. Consequently, cognitive representation is not determined by the "stimulus" alone but always in connection with the organizational and other response dispositions characteristic of the person.

Thus cognitive theory differs from S-R models on several counts. First it neither makes proximal factors the focus of determination nor does it introduce distal factors in disguise. Second, the theory does not derive the organism's coordination to the geographic world from associations between individual stimuli and specific responses. Third, it includes a distal determination in its conception without omitting reference to the mediating processes. It proposes a field-determination which is neither distal nor proximal alone, but includes both these factors as aspects of the *total* field determination.

PERCEPTUAL REPRESENTATION AND PHENOMENAL FIELD IN OTHER THEORIES

We are now prepared to examine the extension of the representational character in phenomenal organization to other areas of cognition. First, it seems apropos to sample the manner in which the configurational tenets have been assimilated by other contemporaries, both with and without S-R leanings. For the Gestaltists the total field, in the widest sense, consists of the geographic environment and the psycho-physical organism within it. Only a portion of this total field is psychologically represented; this is the phenomenal (or psychological) field of direct experience or awareness. It includes distal representation and the phenomenal self. The behavioral environment, or life space, is more inclusive than the phenomenal field, and extends beyond momentary awareness. In turn, the behavioral environment corresponds only to that segment of the wider geographic environment that affects behavior.

Hull (1945) has acknowledged configurational precepts with his concepts of "stimulus patterning" and "afferent interaction." The latter is peripheral in nature, not really involving central processes. According to Hilgard, "the postulate of neural interaction . . . represents an honest recognition that the problem of patterning in perception is a genuine one. The postulate does not solve the problem . . ." (1948, p. 82). It relegates the issue to the nervous system, where every combination of afferent impulses can produce patterns, but the theory fails to specify in what way and under which conditions.

Hebb (1949) grants phenomenal organization insofar as he accepts an unlearned "primitive unity" as the "distinctiveness of figure from background." Such a visual boundary reaction even occurs in rats reared in darkness (p. 20, ff.). Distal representation is thus envisaged. In fact, cues from the distant environment are shown to govern primitive space perception more than nearby objects (p. 92 ff.). Learning and "nonsensory factors," however, are called for in the discrimination of forms, i.e., pattern differentiation as "identifiability." The learning presumably consists of a visual tracing along the figure boundary and a corresponding build-up of neural cell assemblies in the brain. Hebb supports this hypothesis by drawing on von Senden's (1932) case reports of congenitally blind patients. Postoperatively, they seemed to identify shapes through learning. This interpretation has been challenged by Michael Wertheimer (1951), who re-examined von Senden's data for such evidence. It is significant that Hebb's theory, in contrast to a peripheral connectionism, does allow for configurational and central factors of representation.

Gibson (1950, 1951) adopts figure-ground articulation in his phenomenological analysis of the visual world, but derives it from tridimensional "solid" vision. In explaining the linkage between phenomenal space and the distal stimuli, however, he considers it unnecessary to invoke a mediating sensory organization. The latter "becomes gratuitous" because the proximal stimulus distribution and the distal object have properties which are strictly correlated (1950, p. 216, 61 ff.). A specific order of light energy gradients on the retina, rather than the formed image, is coordinated to the surface and depth characteristics of the physical object. Such "ordinal stimulation" is used to account for direct visual form perception as well as the constancy phenomena. This leads to a distinction between the "literal visual world" which is formed by the ordinal stimulation, and the "schematic visual world" which is a centrally mediated representation. The latter consists of the abstractions and meanings which the individual imposes upon his literal perceptions. The schematic view is synonymous with cognitive structuring, which is attained through learning and preverbal conceptualizations. It culminates in verbal symbolism (1951, p. 102 ff.).

Brunswik (1936, 1952), although objecting to the "thing language" in psychology, subscribes to distal representation consisting of patterned cues and "sign Gestalten." Cognition is treated as the functional achievement of attaining objects or final goal effects through perceptually guided behavior. From his ecological studies on perceptual constancy Brunswik formulates a probabilistic approach to the "gross organismic coming-to-terms with the environment." "Cognitive correctness" must be defined in "terms of over-all statistical correlation between classes of variables" instead of single judgment measures (1952, p. 23). A "psychology in terms of objects" is proposed, i.e., an empirical inventory of the kinds of goal effects and the probability of their attainment. Brunswik's system differs from Gestalt theory in that mediation problems are not considered essential, and the details of actual sensory, nervous, or motor phenomena are deliberately disregarded. "One and the same means-object may be represented at different times by very different stimulus configurations. And one and the same goal may be reached equally well by very different kinds of movements and means-object manipulations" (1936, p. 125).

Brunswik (1952) spelled this out by presenting a lens model of molar behavior. The distal stimulus is called the initial focal variable, and the accomplished result the terminal focal variable. The mediation between the two varies according to the circumstances and is therefore "multiple" or "vicarious" or "interchangeable" in terms of a family hierarchy of cues and habits. The cognitive terminal focus can remain substantially invariant. Perception, because of its probability character, becomes a "quasi-rational system" and the lens paradigm a conceptual tool for "the operational definition of meaning" (1952, p. 20 ff.).

Other authors also retain the phenomenal field concept without following Gestalt theory in its specific tenets of mediation. For example, the postulate that "the relationship between the behavioral and the physiological field is crucial" (Koffka, 1935, p. 53) has not been adopted by Lewin. Using the concepts of the life space and psychological field, he formulated his field theory. But he did not include in his system mediating physiological processes, which he considers among "alien factors." Lewin developed principles of behavior in which cognitive representation was objectively dealt with (cf. Chapter 4).

Snygg and Combs (1949) lean heavily on the concept of the phenomenal field and the characteristics of figure-ground organization. However, these authors advance a different conception of life space. It is not an objective construct, as in Gestalt and field theory. The phenomenal field which is posited coincides with subjective awareness. Aside from consciousness, no other psychological events are deemed necessary. The problem of unconscious processes or of a behavioral environment extending beyond awareness is relegated to hypothetical degrees of conscious clarity. Prediction has to be based upon the understanding of the phenomenal field of the behaver, as perceived by him at that moment. This understanding can be accomplished by a method of inference or reconstruction of the behaver's field which must include his "phenomenological needs" and "the projection of the future field." The observer does this by asking typically, "Under what circumstances would I have done that?" Here the subjectivity and uniqueness of the representations in the person's phenomenal field is the paramount material for the psychologist. All influences outside the phenomenal field belong to the realm of an "objectivistic" psychology which is rejected. Reference to the organism as a biological system and to the geographic environment is avoided. The total field ". . . is made up of the individual's meanings of events past, present, and future" (p. 248). The person's behavior is therefore completely determined by his "private" world and conscious frame of reference. This poses the problems of how each phenomenal field can correspond to reality: how it can have access to other persons' private worlds, and how commonality with others is possible. The problems are not solved by postulating an objective geographic or a real behavioral environment that is shared by all human organisms. The answer is given in two parts: (1) There is some connection or relative correspondence between the phenomenal fields of different persons which is based upon commonly shared *experiences* in a given culture or subculture (p. 108). (2) "From a phenomenological point of view 'objective' facts are derived from the phenomenal fields of a number of observers" (p. 249 n.). These facts "represent the least common denominator" of the differing phenomenal fields (l.c.). The individual's world therefore consists of the "meanings" which the "objective facts" have for him. The core concept of this system is the phenomenal self. It "is the only frame of reference which the individual possesses" (p. 79). The maintenance and enhancement of this self is the "prime objective of the individual's existence."

The theory, then, establishes the psychological realities by borrowing from Gestalt and field theory the organizational characteristics of cognitive representation. The role of cognition is markedly tied to an egocentrism of perspective and is placed within a framework of cultural relativism. For the upholder of phenomenological tradition this method would qualify as a rather idiosyncratic version. He may wonder about the procedure of establishing the non-subjective reference of experience (i.e., "objective reality") via a consensus reached through nose counting of phenomenological fields. And he could doubt whether the proposed method of inference is truly phenomenological, since Snygg and Combs reject introspection by the behaver as an instrument of reconstruction. Introspection, they believe, distorts and modifies the phenomenal field, whose "indistinct characteristics . . . cannot be verbalized and therefore cannot be communicated except by inference from their behavior correlates" (p. 36).

MacLeod (1951, 1947) follows more closely in the footsteps of the Gestalt school and Husserl's early theorizing. He extends phenomenological technique to the sphere of social psychology. Psychological phenomenology is not a school or a system but an approach. "A phenomenon is by definition an appearance. Psychological phenomenology is the attempt to suspend biases and to observe and describe faithfully that which appears" (1951, p. 226). "The phenomenological question is simply, 'What is there?,' without regard to Why, Whence, or Wherefore" (1947, p. 194). The method is claimed to be objective. Without previous assumptions it describes what is given in direct experience — things as they are. Moreover, phenomenology is assigned the propaedeutic task of defining psychological data, deciding which are important for theory, and setting the initial problems. From there on, scientific constructs and explanations should be built which may even include physical correlates. This differs from the finality of Snygg and Combs' subjectivistic premises and their procedure of inferential reconstruction. For MacLeod, too, perceptual representation is the point of departure, inasmuch as he accepts the structure of the phenomenal field as given, and the content as including things and persons in relation to the behaver. The relationship is not solely subjective in nature. To speak of the phenomenal world as though it were only a private world tends to create confusion. "There is nothing in phenomenological analysis which requires us to believe that the self is the only, or even the main, point of origin of vectors" (1951, p. 236). The category "mine" characterizes only some of the components of the phenomenal world. "The bulk of the phenomena are not mine at all, any more than they are yours. You and I observe and describe the same object . . . for each of us there are some phenomena of a private or subjective nature but together we live in a public world, a world which contains people as well as things" (p. 237). The danger of confusion, it may be reasoned, arises as follows: One begins with the undisputed fact that ultimately all experience depends upon an organism, a person. Direct (or immediate) experience is therefore only subjective, and objective reference can be established only indirectly or by inference. Consequently, private and immediate experience become synonymous. The phenomenal field has no *directly given* objectivity, i.e., objective contents or reference. Operationists have not made things clearer when most of them label direct experience as private and at the same time admit that all science ultimately rests on this mode of experience (Boring, Pratt, Skinner, Feigl, 1945). MacLeod no less than the Gestaltists insists that the very nature of direct experience itself has subjective *and* objective reference; it already has a discriminable public and private character. To paraphrase MacLeod, there is, for example, an undeniable difference between color and pain. Usually the color is "out there" and a property of an object. It remains there in my absence. It is not a property of me. In contrast, the pinprick on my hand is mine; it is felt in me and not in the pin. I cannot push it out, making it independent of me. Objectivity and subjectivity are phenomenal facts of the same order and validity as the attribute of brightness or of loudness. They are "properties of an organized perceptual field in which the points of reference are selves (subjects) and objects . . ." (1947, p. 200). From a cognitive point of view, the self and the object are part of the same perceptual field. However, the self is an ever-present anchorage point, with the emotional quality of selfhood feelings. Only under particular circumstances (e.g., primitive or early developmental stages, or border states of consciousness) does the phenomenal field lose this dual representational reference of self and not-self.

If one accepts this analysis, serious doubts arise concerning the taboo placed on direct experience as unreliable material for psychology. If the immediately given already contains discriminable objective and subjective characteristics, one cannot call all direct experience private. Moreover, even the "private data" may be accessible and amenable to psychological investigation. To relegate what is private to the realm of the scientifically irrelevant may leave out what is psychologically relevant. A distinguished physicist, widely respected among psychologists, has this to say:

> Simple observation shows that I act in two modes. In my public mode I have an image of myself in the community of my neighbors, . . . in the private mode I feel my inviolable isolation from my fellows. . . . To ignore this difference, or to set up the thesis that the difference is unimportant until it is proved, is opposed to the entire spirit of the operational approach. . . . The operations by which I now deal with the "public for me" are qualitatively different from the operations by which I now deal with the "private for me." . . . Since one aspect of the structure of my universe is the operational difference between mine and thine, I must make my report in a language which recognizes this operational duality. . . . It may be that eventually we shall be able to take account of the operational dichotomy in the universes of each one of us in some simpler way. But until that time, it seems to me that by ignoring the dichotomy psychology is engaging in an unnecessary gamble and is probably riding for a fall. (Bridgman, 1945, p. 282 ff.)

Bridgman asserts further that it depends upon the subject matter of a discipline whether the private and social modes of individual behavior should be recognized and insisted upon; while the social mode may be ignored in physics, this is not the case in psychology.

A methodologically interesting contribution to this question has been made by Wood (1940). He discusses the confusion between direct experience and private experience which has led to a misidentification of phenomenology as introspection. To clarify this situation, Wood introduces the distinction between inspection and introspection. The first is an elementary form of cognition which is immediate. It is a primary experiential activity directed toward its appropriate object and at the same time it is not self-cognizant, i.e., not self-reflective. Introspection is a secondary subsequent cognitive activity in which the primary cognitions are reflected upon. The difference between these two cognitive modes may be illustrated as follows. Perceiving the action of a person is inspectional cognition, since I do *not include my seeing* in the perceptual act: the-action-of-a-person-seen. Introspection enters when I reflectively perceive the total situation as myself-in-the-act-of-perceiving-the-action-of-the-other-person. Wood claims that inspectional data correspond to direct experiences, are accessible to all, and are essentially the same data with which phenomenology and Gestalt psychology deal in perception. Introspectional data are private, inasmuch as they are accessible only to the introspectionist himself. In this sense, introspectional cognition is a constructional knowledge "of a peculiar class of objects, namely, those belonging to the continuum of the introspectionist's personal experience. Introspection is dependent upon inspection, for, although introspection involves and presupposes inspection, the converse is not true — inspection may or may not be introspective in character" (p. 227). In this connection it is noteworthy that Brunswik, notwithstanding his objectivistic functionalism, concedes this role to direct experience. "It must be stressed, however, that act-psychology and phenomenology are important, and perhaps indispensable preliminaries to molar research of the objective kind" (1949, p. 59).

MANIFEST ORGANIZATION AND COGNITION

Employing the principles of phenomenal unit formation as the basis of distal representation, Gestalt theorists and those of similar orientation have developed the special characteristics of the phenomenal field. This field always represents something — some content in direct experience. Since direct experience has molar and relational characteristics, we may speak of the "manifest organization" of the phenomenal field. Manifest organization is awareness of behavioral objects and of their relatedness to each other and the relation between their properties. Without manifest organization, awareness would be a mere sequence of experienced events. In the words of Koffka, ". . . there would be no forces. One billiard ball would run, come in contact with another, stop, and the other would begin to roll. Two trains would collide, leave the tracks, and cars turn turtle and become wrecked; another mere sequence" (1935, p. 383). This is the kind of world that many S-R theorists, like Hume before them, posit. The contiguity of events in space and time leads to the mental habit of uniting these sequences as cause and effect. The experienced "causal union" becomes felt necessity only if the same *contiguities always* occur.

Phenomenal causality, however, according to the Gestalt laws of unit formation, is part of the manifest organization through which the why and how of objects affecting each other can be directly experienced. Duncker (1945) and Köhler (1947) have made searching examinations of the characteristics of phenomenal causality. Both discuss its intelligible aspects and its cognitive relations with the objective environment via the behavior space. There is first the functional aspect of objects, which is learned and grasped through action, e.g., the function of the hammer in driving the nail. The experienced functions are not arbitrary but "ground consequence structures" in which the properties of the object are intrinsically related to its use and effect. There is, further, an intelligible connection in the experienced correspondence between the function of objects and their material or their form. The *material* fits to the functional "valence" of things according to whether the material is light, heavy, soft, hard, flexible, or immobile. Correspondence of *form* and function is seen in the rolling of a ball, the notches on a key matching the casing of the keyhole, and the shortest detour conforming to the contour of an obstacle. Intelligible connections are also seen in the correspondence between the *phenomenal properties* of cause and effect. Although these phenomenal properties overlap, they may be enumerated.

I. *Positional correspondence,* as in (a) spatial proximity: the effect is perceived close to the place of its origin, e.g., the sound is heard *where* the bell rings; or (b) temporal proximity: the effect follows closely its antecedent, e.g., the landscape is illuminated after the flash of lightning.

> II. *Qualitative correspondence,* as in (a) congruency: the effect is in harmony with its cause, e.g., the trace in the sand fits the form of the object, as in footprints, or the sound of knocks corresponds with the rhythm of the knocking; or (b) similarity or identity of content: one event passes from one medium to another but retains its character, e.g., the heat of the stove becomes the heat of the room, the wetness of the rain the wetness of the pavement, the blackness of the pencil point the blackness on the paper.
>
> Phenomenal causality therefore appears to be dependent on the objective contiguity of events. Hereby results a phenomenal coincidence which is not accidental or arbitrary. The factors of proximity, identity, similarity, and good continuation as they are inherent in the particular contents of events become codeterminants of the phenomenal coincidence, so that a *selection* occurs. Cause and effect depend upon positional and qualitative correspondence; in every case of direct causal experience this *functional connection stands out perceptually* from *all other conceivable simultaneous relations and events.*

This analysis is, in effect, the reverse of Hume's position. Because in experienced causality the time and place of the cause "coincide" phenomenally with the time and place of the effect, there is sometimes a tendency to link causally those events which have no causal connection save contiguity. The mere coincidence of a door slamming behind a person in a dark hallway and the hallway light going on at the same time can be to him a compelling impression of causation *without* intelligibility. The factor of proximity is preserved, but the *functional* connection between cause and effect is phenomenally absent. In contrast, it is ever-present in our actions of manipulating objects where we are the causes of the change in the object, e.g., when we break a stick. It is therefore not surprising that Köhler (1947) applied the term "insight" to the manifest organization of phenomenal behavior and also of the ego-field structure in general. Both can have the characteristics of "awareness of determination." The cognitive aspects of manifest organization *can* include the awareness of the dynamics of our own action: not only do we act, but we know why we act, and why our acts have effects.

A commonplace instance of phenomenal causality in behavior is action in terms of means-end relationships and detour-paths to goals, i.e., "this leads to that." Ever since the experiments by Köhler, Yerkes, and others with chickens, chimpanzees, and children, the principle of detour has been widely accepted. We have here an affirmation of Tolman's position that molarity, goal directedness, and cognition are necessarily interdependent. From any segmental point of view, each single movement away from the goal would actually lead away from it unless it is phenomenally related to the total detour path as part of the whole. It is in this context that Tolman and Brunswik (1935) analyzed the "causal texture of the environment" and extended the concept of representation. In the organism's environment certain objects may function as the *local representatives* of other objects; the latter then are considered the entities represented. The function of local representation in the causal texture of the surroundings has two subvarieties. The first is the case of tools or paths which are "behaved to" only because of their role as means-objects. They function "as 'local representatives' whereby to reach or avoid such and such more ultimate, 'represented' positive or negative goals" (1935, p. 45). The second is the selective perceptual response to proximal stimuli ("light-wave bundles and the like") as cues or signs, i.e., local representatives, of the distal objects or events. The latter may also function as representatives for still other objects or events and may then come to serve as a cue or "sign Gestalt" for a causal relationship. For example, "the retinal effects on the child's eyes serve as local representatives, signs, for the facial expression of (another) individual. And this facial expression, as a local representative, serves in its turn as a sign . . . for the will, good or bad, of the other individual" (1935, p. 51). This approach essentially parallels the Gestalt tenets of phenomenal causality but introduces one important variation. The parallel lies in the assertion that in the environment there exist causal couplings in which different events are regularly dependent upon each other; that because of this, organisms come to accept one event as the local representative of another event; that they thereby come to "behave" appropriately to the more distant, "represented" situation or event. The variation lies in the assertion that the causal connections thus represented always have some degree of equivocality, so that a *probability* factor is attached to the local representative. The organism approaches the environment with certain hypotheses and expectancies about the most probable causal relations and constancies.

APPARENT BEHAVIOR: COGNITIVE STRUCTURING IN SOCIAL PERCEPTION

During the last decade the social implications of cognitive structuring in phenomenal causality have been investigated by several authors. Michotte (1950) performed a series of experiments using simple geometric figures, e.g., two colored rectangles moving towards or away from each other, under varying conditions of speed, direction, and distance. The subjects were presented with a variety of combinations:

1. *Object A moves toward and reaches B, which is stationary.* When *A* moves rapidly, the impression is that it strikes *B;* when *A* moves slowly, it appears to gently join, touch, or unite with *B.*
2. *After A reaches B and stops, B begins to move.* If *A's* approach is more rapid than *B's* subsequent motion, *B's* "moving away" seems to be caused by *A,* sometimes interpreted as a reaction of anger. If, however, *A's* approach is less rapid than *B's* subsequent motion, then *B's* movement appears autonomous, only released by the contact with *A,* and gives the impression of flight, sometimes interpreted as a fear reaction of *B.* This occurs when the duration of the contact is short. When the contact lasts longer, "the impression produced is often interpreted in the sense of a momentary agreement between two accomplices who meet, but which gives place to a disagreement, followed by separation."
3. *Object B begins to move after being touched by A and the two continue to move together at the same speed and in a common direction.* If there is a short pause after contact, the impression is one of friendly "getting together," the unity being accentuated by the similarity of the speed and direction of the subsequent movement — the "law of common fate." If there is *no* pause, and *A* approaches *B* slowly, with rapid acceleration during contact, *A* appears to carry off *B* by brute force. If, under the same conditions, the direction of *A's* approach is reversed during contact, *A* may appear to snatch *B* back with him ("kidnapping").

Michotte considers these impressions to be Gestalten, namely, specific perceptual kinetic structures related to the action of successive stimuli on the eye. The lessening of the distance between the retinal images of the two objects causes the impression of approach, but this is much more than a simple response to diminution of distance. The perception is selective because in a complex visual stimulus field, the lessening of distance between two objects is necessarily accompanied by increasing distance between the moving object and *others:*

> "Yet the impression of 'going toward *B*' is never accompanied by the simultaneous impression of 'going away' from these others. . . . There seems to be some sort of incompatibility between the simultaneous presence of impressions of 'going toward' and 'going away from' in the case of the same object. It is something analogous to the distinction between 'figure' and 'background,' where contour belongs exclusively to 'figure'." (1950, p. 120)

According to Michotte, the organizational factors of perception are autochthonous unit formations which vary in strict correspondence with changes in the stimulus conditions, e.g., rate of movement, direction, and distance. Thus the category of causation is believed to have the same immediacy as shape or figure-ground relations. Perceiving movements as properties of figural units, we refer them to definite origins and so connect consequences with antecedents in terms of a constant object passing through and producing change. Specific sequences of motion are thus perceived as "functional connections." These experienced causal contexts are also described as *felt* connections by Michotte, who thereby bridges the traditional separation between cognition and emotion. The perceptual response includes similar somesthetic and motor reactions in the "affected" observer, as in the experienced expressive movements of other persons. Studies pointing in a similar direction, linking cognition and emotion, have recently been made by Werner and Wapner (1952). The interrelation between the tonic state of body equilibrium and perceptual forces in the visual field has been experimentally explored, leading to a sensory-tonic and viscerotonic theory of perception. The visual dynamics of faces or triangles pointing right or left when presented in a dark room affected the sensory-tonic equilibrium correspondingly. A right-pointing profile or triangle was measurably displaced to the right of the median plane of the observer, etc. This approach opens an important new avenue to establishing empirically the intraorganismic processes attending the perception of "physiognomic" or "demand qualities."

All such findings have far-reaching applications to the perception of social action. Michotte's subjects spontaneously used not only the language of causation but also spoke in terms of human emotions and attitudes. They described the inanimate figures in such phrases as "It is as though *B* was afraid when *A* approached and went off . . ., *A* joins *B,* then they fall out, have a quarrel, and *B* goes off by himself," and so on.

Heider and Simmel (1944) independently arrived at very similar results after showing a movie to several hundred college students. The movie presented different geometric figures, e.g., a small triangle, a large triangle, a circle, and a rectangle in varying combinations and movement sequences of changing speed and direction. The subjects' verbal descriptions showed a striking preponderance of animistic responses, describing the figures not only as human but also in terms of interpersonal relationships such as rivalry between two men about the girl, or family conflicts among mother, father, and child. The second overwhelming tendency was to structure the movement sequences into units which comprised a meaningful action in which the three forms interacted in a causal and motivated fashion, e.g., chasing, attacking, fleeing, and so on. These actions also became related to the fourth figure, the large rectangle, part of which opened and closed during the film. Over ninety percent of the subjects described this rectangle as a house which the large triangle angrily destroyed at the end of the story. Heider's analysis demonstrated once more that the relations of temporal succession, spatial proximity, good continuation are decisive stimulus conditions for the causal and interpersonal perceptions involved.

In all of these experiments the influence of past experience upon the perceptual reactions of the observers may appear obvious. However, if the explanation of directly apprehended phenomenal causality were to be derived from past experience alone, it would have to account for the lawful stimulus response correlations obtained by Michotte and Heider. Of course, the reported "animation" points to the similarity between the movement forms of the abstract figures and the previously experienced movement forms of concrete persons. Yet to derive animation from this similarity does not answer the primary problem: What are the *original* and relevant stimulus conditions for the perception of human action, and how do these phenomenal unit formations arise which are germane to perceiving *others?* Heider (1944), Asch (1952) and MacLeod (1947) have each dealt with these questions in rather similar ways.

Under everyday conditions objects are not usually experienced as agents which *actively* cause events (except in animism). Objects are related to each other in causal *determination.* The cigarette scorching the tabletop is not seen as the cause of damage, but its burning (and being left unattended) is so seen. The laws of unit formation operate differently when we perceive the behavior of objects and when we perceive the behavior of persons. In human action, we perceive the person as the causal agent or source from which the action issues. The person and his action perceptually form a specific causal unit in which the person is the author of the action and its effect. The perceived direction and attainment of his action depends upon an organization of movements according to spatio-temporal laws of proximity, similarity, and good continuation in varying degrees. Inasmuch as the act and its accomplishment fit the perceived direction, as, for example, in attack or in pointing to an object, we also perceive a congruence between the motive, goal, or intention of the behaver and the course and terminus of his activity. These principles apply just as much to misperception as to the correct perception of human actions.

Heider (1944), following the leads of W. Stern (1938) and Werner (1940), analyzes different levels and modes of perceived personal causation ranging from the more primitive (magic) to the more complex. Different cognitive structures are involved in such cases as: (1) The primitively experienced union of the doer and the deed — the person who inadvertently steps on my foot *is* the cause of my pain. (2) The more differentiated causal relationship between the doer and the deed — the dentist who unintentionally causes me pain is not perceived as the cause of my pain. This also has significance for the relationship between frustration and aggression. Aggression is more readily evoked and more intense if the frustration is attributed to a person rather than to impersonal causes. The application of these observations to other interpersonal relations, such as forming impressions of personality, has a wider orbit than can be discussed here (Asch, 1946; Mensch and Wishner, 1947; Chapter 17, this work).

The general maxim should be stressed that by virtue of these specified unit formations the "other" comes to be represented as a causal entity in the phenomenal field of the person. Thus "the paramount fact about human interactions is that they are happenings that are *psychologically represented* in *each* of the participants" (Asch, 1950, p. 142). In the person-to-object relation, representation is *one*-sided, whereas in the person-to-person relation, it takes place on *both* sides and is interdependent. Psychological interaction between person and person occurs "between entities both of which are systems which perceive, feel, and understand." This leads to the concept of the "mutually shared field," which includes both the common behavioral environment and the psychological perspectives of the interacting behavers as the objective sphere of interaction. Even the most simple behavioral interplay, let alone cooperative acts, requires a reciprocal *representation* in which "different and complementry actions are executed simultaneously with reference to each other" (Asch, 1952, pp. 163, 175). In his analysis, MacLeod (1947) points to the phenomenological fact that a person has a unique status among other objects of the psychological field, that by being apprehended as an "other self," he is more or less "like me" and is at the same time apprehendable as an "object" which can generate its *own* purposes. Our dislike of a person will be felt as caused by, not as causing, the perceived purposes or properties of the person.

The fact that these principles are still in need of empirical elaboration is recognized by MacLeod, who urges the use of improved methods of observation, especially of everyday life situations. A fruitful approach in this direction may be the dividing of the behavior continuum according to criteria derived from the principles of unit formation in social perception, a method proposed by Wright and Barker (1951). Action units are defined as molar behavior episodes oriented toward a molar environment, i.e., towards situations. The molecular aspect of actions, such as steps in walking, are defined as subordinate and mediating part processes. Among other criteria, behavior episodes have the following molar and whole-part characteristics: invariance of direction; relative inclusiveness or size which varies with the subgoals and subunits; the span of the person's own behavior perspective in his current goal pursuit. These criteria are further refined with regard to relative potency of the phases of an episode, its transitive and consummatory stages, the degree of completion or incompleteness of episodes and their overlapping. Related molar criteria for perceived behavior units have been developed by Krech and Crutchfield (1948, pp. 30 ff.).

All the foregoing discussions of social perception share this in common: they deal with the stimulus as a cognitive representatior. Moreover, this representation can be conceived as causing behavior, and the recognition of this can lead to predictions. Here again cognitive and behavior theorists are divided on a fundamental issue of method. Cognitive theorists not only recognize the relevance of the behaver's phenomenal field for all psychological theory, they also consider it a cause, i.e., an intervening variable or an explanatory construct of his actions under certain conditions. Therefore psychological understanding and explanation can coincide for cognition theorists, who in such cases would invoke causal principles for their conceptualization. For the S-R advocates with strong operational leanings, such as Kendler, this leads to the danger of not clearly distinguishing "the requirements of scientific explanation from those attributes of propositions which are capable of instigating in some a feeling of 'understanding' (the 'a-ha' phenomenon)" (1952, p. 274). We must conceive the intervening variables "as *economical devices* to order experimental variables in relation to the dependent variables." Psychologists are criticized for going beyond the operational meaning of the intervening variables by reading into them surplus, particularly causal, meanings. "Most, if not all psychologists were confronted with the question 'Why?' prior to their exposure to scientific training" and their "pre-scientific explanatory techniques have lingered on and been confused with scientific explanation" (p. 274). Apparently causal understanding has been placed outside the realm of science and into that of "personal cogitation" and "intuitively satisfying explanation." Ritchie (1953) satirically criticizes Kendler for this proposed "circumnavigation of cognition." We may ask whether the following instance can be conceptualized as an "intelligible" causation from Kendler's point of view. When an observer sees a motorist avoid a child on a highway, he will experience the same causal connections as the driver who perceives the child and avoids it. Here is an experience which causes behavior and is perceived as such by an observer. For Kendler, causal explanation is to be substituted by either formulating empirical relationships with predictive capacity or constructs of deductive capacity. This rationale is a direct outgrowth of Feigl's proposal that "low-grade" explanations in science be replaced by "high-grade" explanations, even though it is admitted that psychology has yet achieved only a few of the "low-grade" (Kendler, 1952). Feigl (1951) considers "high-grade" explanation that which relegates macroscopic events to underlying microphysico-chemical events. The problem here is: How can the logical positivist bridge the gap between psychology and physics without interposing the organism as a system? Failure to recognize the primacy of functional and structural organization in an organism inevitably leads to an uncalled-for reductionism which skips the entire dimension of biological data. Rather than prematurely delegating causal thinking in science to the realm of the subjective, prescientific, and "low-grade," there exists the need to account scientifically for the very phenomenon or causal thought. This need has been accepted by the cognitive theorists. Duncker (1945) has analyzed the conditions for partial and total insight in physical science. He shows that there are ultimate irreducible facts which are not intelligible. However, the formulation of laws which *relate* these facts to each other often do qualify as intelligible. They must even have *logical* consistency, particularly if they are couched in the language of the hypothetico-deductive method so ardently and rightly advocated by Feigl (1951).

The assertion that truth, understandability, and logic are reducible to mere intersubjective consensus because science must be "public" has been challenged by Bridgeman (1940).

> The process that I want to call scientific is a process that involves the continual apprehension of meaning . . . accompanied by a running act of checking. . . . This checking and judging and accepting, that together constitute understanding, are done by me and can be done for me by no one else. They are as private as my toothache and without them science is dead. . . . When I follow the marks on paper that constitute a treatise on symbolic logic, I am not willing to call my activity scientific unless I can say at each stage of the process "I understand." . . . (1940, pp. 41, 46)

WHAT IS A RESPONSE COGNITIVELY CONSIDERED?

It seems difficult to define the relevant stimuli for social perception and thinking in physical terms alone. They have meaning. S-R theorists typically refer present meanings to past associations. The manner in which cognition theorists have dealt with the relationship between response and past learning differs from S-R conceptions. Both schools agree that present responses can always be influenced by past experience. They disagree on *how* the past learning occurred and *how* it affects the current response. To begin with, cognitive theorists have not separated stimulus and response in the traditional way. The stimulus is phenomenally represented through configurational responses to the *relations* between proximal stimuli. The response is therefore already a process of spontaneous organization. In the case of learning the response is also viewed as *relational* rather than specific. For this reason is it advisable to examine first the treatment of the response in the pertinent areas of learning and problem solving.

There has been a tendency to exaggerate the reservations, voiced by Gestalt theorists, about the role of past experience. As Wertheimer (1945) emphasized, the issue is not present structuring versus past learning, but rather *how* learning takes place, whether primarily through associative couplings or through organizing the relationships inherent in the content. It is held by cognitive theorists that genuine learning occurs through structuring and restructuring of the psychological field. In this respect, learning and problem solving have cognitive processes in common.

Thus, maze learning by animal or man is not viewed as a stimulus response linkage in the sense of peripheral receptor-effector connections. Rather, the learning is viewed as the acquisition of central representations of layout, place relations, and directions, viz., reaching the goal by swimming after having learned to reach it by running, or reaching it after severance of the kinesthetic tracts. The experimental fact that new adequate goal paths can be selected when those learned previously are blocked led Tolman to propose the "cognitive map" as an intervening variable. Once acquired, such cognitive structures (sign Gestalten, field expectancies, etc.) are available for the organism. They can be activated when future needs arise. These representations can subsequently be reorganized when detours are demanded.

> Light is thrown on the organizing processes in learning and problem solving by the experiments of Köhler and others with primates and children (Köhler, 1925; Alpert, 1928). A gap between the reach of the arm and the food lure is filled. A stick is "seen out" from its configurational context and its current properties are "recentered" into "rake-in and arm-extension" properties. One trial alone is sufficient for learning. Also transfer to other objects with gap-filling and rake-like qualities occurs spontaneously when a stick is not available. This militates against the interpretation that specific associations are formed in the original experience. The achievement is rather rooted in the learning of functional connections between object properties and action. The understanding of such causal structures involves a certain cognitive generality. The famous double-stick solution of Sultan occurred later with other sticks of new diameters. The transfer consists in the carry-over of the experienced joinability of hollow sticks which can be wedged into each other. Köhler's theory of a predominantly visual structuring has been supplemented by Birch. He showed that chimpanzees without any previous stick experience could not solve the raking-in problem within thirty minutes. After having played with sticks they achieved solution within twenty seconds. The interposed experience was carefully controlled. The animals were observed to use the sticks as extensions of the arm for poking or making contact from a distance, but they never used them to sweep objects into reach (Birch, 1945, p. 375). The use of the stick for raking is therefore a "new" pattern of response in which the preceding stick experiences are *reorganized.* This stick experience may be necessary action-wise, but again does not consist of specific, "contiguously learned items of experience" (p. 381).

Concerning the use of previous knowledge in problem solving by human subjects, the issue is the following: Can insight be reduced to a response which repeats an identical or partially identical element of past experience, or repeats identical reactions to equivalent stimuli, as Boring (1946) and others claim? In the first place, it has been shown that past experience often hinders solution in problem solving, to wit, Duncker's (1945) investigation of "functional fixedness," and Maier's (1930, 1931, 1933) research on "direction" in reasoning. Second, Maier (1945), in a variation of his experiments on hatrack construction, has challenged specifically the doctrine of identical elements and of equivalence: reproductive thinking, defined as a transfer of previous knowledge, is only operative when the identity or equivalence is discovered phenomenally. "Objective identity is wholly inadequate." In productive thinking not only the present but also past experiences can be changed and restructured. Stimulus properties which have not appeared before may appear under the stress of a problem situation. According to Maier, this militates against the traditional assumption that only the perception of the present thing may change during learning or reasoning; that the memories of this perception, however, cannot be modified and reorganized because they are static.

With his experimental design Maier did not probe beyond reproductive thinking in which subjects had to discover an identity between an item in their past life experience, e.g., a hatrack, and a presently available item, e.g., a clamp on a pole. Nor did he investigate productive thinking beyond the reorganization of the memory for a specific item into a different familiar item; e.g., a clamp which first had been used to join two poles for a supporting column later had to be recentered into a hook, creating a hatrack. Maier's definition of productive thinking could be widened to include the creation of a new aspect which is not contained in a specific past experience. Wertheimer's (1945) analysis has strongly suggested such genuine creative structuring in the instances of young Gauss and Galileo.

Restructuring may also transcend the discovery of a specific familiar item. In the solution of social problems, for example, a re-examination of the current goal or intention may be required. This specific goal is taken out of the present context in search for the *broader underlying* intention. The discovery of the "true" nature of the goal can produce a novel organization of the means for its attainment. Wertheimer (1945) devoted a chapter to the conflict which two boys encountered in playing badminton because the older one was so superior to the younger that the latter refused to continue the game. The solution: the older discovers that when both join in the common goal of keeping the bird up in the air as long as possible the game can remain enjoyable for both. The boy re-examined the need for his present competitive goal and found that the underlying intention or true goal of the game was to have fun together. This leads to the recentering from "playing to win" to "playing cooperatively."

In another type of productive thinking an unfamiliar problem may be reorganized in a novel way for the utilization of past learning. This is the case when a child has once grasped the structural principles for determining the area of a rectangle. He may then find the area of a parallelogram by discovering that there is "too much on one side" and "something missing on the other." He reorganizes his percept and "sees out" the superfluous triangle on the one side and fits it to the other side of the parallelogram in such a way that a rectangle is created.

Transfer effects in terms of relational principles have been experimentally demonstrated in Katona's (1940) experiments on card tricks. The group which learned by understanding could subsequently solve both the original and new card tricks, while the group that learned by drill showed poorer retention on the old tasks and failure on the new ones. An important feature of this, as well as many other studies of this kind, is that even the S's who demonstrated "understanding" were not readily able to verbalize the organizational principle involved. Experiments by Szekely (1940, 1950a,b) on the creative restructuring of latent knowledge specifically suggest the presence of implicit relational thinking. A similar conclusion was drawn by Leeper (1949, 1951) in reviewing the work on determining tendencies and concept formation. He indicates that there is an intimate connection between the cognitive structures of mental set, imageless thought, transfer in problem solving, and certain processes of generalization. A common denominator of these processes is that they usually are outside of direct experience and consequently unverbalized.

SILENT ORGANIZATION: A COGNITIVE DETERMINANT OF RESPONSE

The inference that only certain aspects of the cognitive organization may be "manifest," i.e., in awareness, poses a new problem. Do "nonconscious" processes have representational characteristics, and if so, how? There are numerous instances that indicate the functional dependence of the phenomenal field on a frame of reference which transcends it in space or time. In the case of the perceived upright, for example, Witkin (1949) has demonstrated the influence of spatial framework which is not part of direct experience. Likewise, the speed and direction of perceived motion may be determined by a spatial frame of reference which is beyond awareness (Brown, 1931; Duncker, 1929). The classical studies on successive judgments of weights have established the effect of a temporarily operative frame of reference not manifest in the subject's experience (Wever and Zener, 1928). Sherif (1936) made these findings the point of departure for his well-known theory of social norms. Sherif and Cantril (1947) reviewed the experimental evidence for the effect of such frameworks on psycho-physical judgments in various modalities. This evidence was utilized in developing a psychology of attitudes originating in perception and governed by frames of reference. Recently Helson (1947, 1948) has attempted to subsume a wide range of similar phenomena under the rubric of "adaptation level" as a unitary explanation.

The Gestaltists have introduced the concept of "silent organization" to account for factors outside of awareness which influence behavior. Among such factors, Koffka has extensively discussed the functional effect of the spatial framework upon body posture in action and locomotion, and upon the localization of the self (1935, p. 216 ff., p. 384 ff.). In the case of silent organization the individual's conscious judgment or performance is thought to be dependent upon a nonconscious frame of reference similar to the way in which a figure depends on the unfocused background. Applying the laws of whole-part relationship here as well, Gestaltists propose that the phenomenal field can become a functional part of the wider process of silent organization as a whole.

Silent organization will be used here in a broader sense. It is suggested that various problem areas may be dealt with by using silent organization as an explanatory construct. This construct would cover the formation of frames of reference and mental sets, the development of both generalizations and concepts in their preverbal form, and the carrying over of implicit organizing principles from one problem-solving situation to another. The common denominator of these silent organizations is that they are not tied to specific individual contents but operate as cognitive structures which have the character of "typifying schemata," guiding orientation and action. Head (1920), Bartlett (1932), Lashley (1951), Piaget (1937, 1952), and others have employed the term "schema" to designate essentially the same integrative processes. If we accept this concept, it follows that silent organization has representational characteristics. This representation does not consist of specific phenomenal objects. It is a trans-phenomenal organization which relates spatial or temporal units of the behavioral field into an integrated structure — the schema. Likewise, relations between ideational or action units of behavior can attain schematic representation.

Such schemata can exercise a striking influence on individual contents of perception and thought. Two experiments, among many others, may serve as illustrations. In Luchins' (1942) study, the S's had to solve the task of obtaining a certain quantity of water with the use of three containers of different capacities. For example, the S is told: you have jars of 18-, 43-, and 10-quart capacity; get 5 quarts. After having solved several such problems, most S's developed a set to subtract from one of the larger containers. The last problem was: given 5-, 15-, and 20-quart jars, get 5 quarts. Certain S's did not realize the direct availability of the 5-quart jar. Instead they adhered to their stereotyped procedure of subtraction. They measured 15-5-5 *or* 20-5-5-5 *or* 20-15. It is not only noteworthy that their schema was more general (subtraction) than a specifically fixed sequence, but also that it literally prevented the S's from *cognizing* that 5 quarts is 5 quarts even though they were able to perceive the 5 and use it for their schematic pattern of solution.

An illuminating case in which schemata of everyday life are tapped is the experiment by Bruner and Postman (1949). The S's were tachistoscopically presented with playing cards, some of which had color and suit reversed, e.g., clubs red and hearts black. The recognition threshold for the incongruous cards was significantly higher than for ordinary cards. The incongruous cards also required a greater number of prerecognition trials. However, when a subject had once recognized an incongruous card, the threshold for later incongruous cards was substantially reduced. Once the possibility of suit-color reversal was recognized, a shift in cognitive structure took place. We have here an instance of the strength of a schema and its reorganization under the influence of changed information. Bruner, without using the term "schema," speaks here of a "hypothesis" which "is best likened to such terms as determining tendency, set, Aufgabe, predisposition. It may be regarded as a highly generalized state of readiness to respond selectively to classes of events in the environment" (1951, p. 125). Again, "a specific hypothesis is not an isolated expectancy . . . but rather relates to more integrated systems of beliefs or expectancy about environmental events in general" (l.c., p. 127). Bruner emphasizes that such hypotheses are dependent upon a "cognitive map." Tolman (1951), in his latest formulations, has extended the map into a cognitive matrix which is a larger system of hypotheses, similar to what we have here called schemata.

The attempt to subsume a whole range of psychological phenomena under the construct of silent organization as a system of schemata may be open to criticism. It may be said that what is gained in the inclusiveness of the construct is outweighed by its vagueness. Of course, the various types of schema building are still in need of empirical clarification. The core value of this conception, however, is that it prepares the ground for a broader unitary theory. A starting point in this direction would be the proposition that the phenomenal field and silent organization share representational processes which can also extend into memory function itself. This is recognized in one way or another by contemporary theorists.

Tolman (1951) contends that perception includes more than what is merely *present* to the senses; it often encompasses "familiar objects on the wall behind my back . . . For I might well be found to be immediately ready to behave to all these other objects." Moreover, the "behavior space" contains not only objects but also their spatial and temporal relations and directions. Even mechanical, aesthetic, and language dimensions may be present. Inasmuch as any of these govern immediate action, they are " 'perceived,' whether or not (in introspective terms) the actor is then and there consciously aware of them" (1951, p. 296 ff.). Tolman also speaks of acquired "typed (not concrete) images of objects . . . conceived to be arranged along functionally defined 'generalization dimensions' " (pp. 290 ff.). These would correspond to the ideational schemata under discussion.

In the last analysis all these formulations point to a need for linking the psychological or behavioral descriptions to physiology of brain processes and traces. The postulate of schemata as silent cognitive representations strongly suggests that certain relationships between the environmental field and the trace system must be preserved. It is not necessary here to speculate on the ultimate nature of this relationship, whether it is, for example, isomorphic. It is important to realize, however, that a mosaic and connectionist theory of the physiological mechanisms involved seems least adequate. Von Bertalanffy, who refers to the work of von Holst, Lashley, Krech, Goldstein, and others, concludes: "The fact that only one experience is in the focus of consciousness at a time seems to indicate that its physiological correlate extends over the whole 'brain field.' If the mosaic theory were correct, obviously any number of excitations and corresponding experiences could be coexistent . . . If, however, form perception is a system process, dynamically ordered and extended over larger brain areas, the aftereffect of excitation will consist not in leaving traits [traces] in individual cells, but in an alteration of the brain field as a whole" (1951, p. 29). In a similar vein, Koffka maintained: "there is no break between the conscious and the unconscious, rather have part-processes of total strains of events the additional property of being conscious" (1928, p. 68). Antedating his later distinction between manifest and silent organization, he states: "Conscious processes, dependent upon unconscious conditions, arise out of unconscious processes, terminate in such, and leave behind new unconscious conditions" (p. 68). Various theories of trace systems have similar implications, namely, that traces are organized or related and not aggregates of elements. Emphasis upon such physiological mediation is urged upon us precisely because silent organization poses the problem of how a transphenomenal representation codifies certain external as well as past events. In this area, the intriguing formulations of Bateson (1951) have been referred to earlier.

BODY IMAGE AND EGO AS COGNITIVE ORGANIZATION

We have thus far considered chiefly the cognitive relationship between external and internal events. In dealing with frames of reference and schemata which mediate response, an important internal framework has not been mentioned. The *ego* is not only cognitively represented in the phenomenal field, but it also exerts an influence in terms of silent organization. The Gestaltists, however unsystematically, have treated the ego in several ways. First, the ego is included in the phenomenal field as a part or a subsystem. Thus it becomes a perceptual object of a special kind. It is that part of the behavioral world which is "in between the front and the behind." It is therefore a system of spatial coordinates which operate as an experienced reference point in the perception of direction. In that sense the ego differs from other perceptual objects because there is always a "mine" relationship involved, e.g., *my* hand; to the left or right of *me;* up and down, etc. This is an aspect of the "personal dimensions of space," as William Stern called them (1938, p. 95).

Second, the ego functions as a silent organization in being a schema of orientation in the behavior space for localizing parts of our bodies, or for localizing our bodies as a whole in surrounding space. The body image as analyzed by Head (1920) and further elaborated by Koffka (1935) and Schilder (1935) is an instructive paradigm of cognitive representation without awareness. They arrived at similar conclusions. The body image constitutes a *central* integration of cues and traces from diverse sense organs into a schema, a plastic model of ourselves. In particular, Koffka adopted Head's view as to the importance of the memory trace system, which forms a silent background or frame of reference in the conscious perception of body posture and position. According to Head it is not the image, visual or motor, which is the standard by which postural changes are appreciated.

> Every recognizable change enters into consciousness already charged with its relation to something that has gone before. . . . Just as on a taximeter the measured distance is presented to us already translated into shillings and pence, so the final product of spatial changes rises into consciousness as a measured postural appreciation. For this standard against which all subsequent changes of posture are measured before they enter consciousness, we have proposed the word "schema." (1920, pp. 60 ff., pp. 723 ff.)

Both Head and Koffka, therefore, made a distinction between process and trace, i.e., the schema *building* and the schema *built*. The latter is the unconscious standard in the form of traces representing the model of the body against which the current process of postural change is appreciated.

We cannot pursue here the finer elaboration of the postural and surface models of the body. Suffice it to say that both are interconnected and related to the ego. The surface model develops into a more constant and value-laden representation of the body. In this connection, it should not be overlooked that Head's theory stems from extensive empirical research in neurology. Koffka made these functional schemata, which are models of ourselves, the starting point for his own theory of a permanent ego substratum or ego system (1935, pp. 516 ff.). An organized trace system develops which integrates ego-related experiences (i.e., "mine" experiences). Only the experiences which refer to the ego or ego system acquire self-character. They form the subjective pole of memory as against the objective pole of an environmental trace system. The so-called "ego-trace column" is the autobiographic frame of reference which temporally localizes every experience as to present, past, and future. It preserves the continuity of the "I" experience, and makes it possible to register cognitively the changes which I, myself, undergo, as well as the changing events in *my* life space. This ego frame of reference can operate silently and also can become represented in awareness, e.g., when I represent to myself an event in my past.

The difference between the environmental and the autobiographic trace systems finds remarkable support in the studies by Claparède and his co-workers (1911) on the disturbed ego reference in the memory of patients with Korsakow syndrome. These patients show *no recognition* of persons or recent events but have functional retention of environmental happenings. They know their way around in the hospital and are well oriented as to the function and rank of the hospital personnel. When faced with their doctor such patients usually fail to recognize him, yet they will be able to "guess" his name from a list of names. Both Koffka and Claparède conclude that recognition requires the cognitive reference of "me-ness," i.e., the experience of "familiar to me."

Other aspects of the ego substratum as seen by Koffka are the motivational content and direction in self-referent behavior and the "executive control" of "my" actions. These are exemplified in the emotions of self-interest and in the experiences of wishing, willing, intending. Here again only certain experiences and emotions are referred to the self. The support for this construct of ego structure is found, for example, in experiments on level of aspiration, on the Zeigarnik effect and on the completion of interrupted tasks. All these findings are taken to indicate tensions which originate dynamically in the ego system and often belong to the dimension of success and failure feelings. Such tensions are therefore distinct from neutral, environmental forces. In this context, attitudes become forces issuing from the ego and they can affect the phenomenal field.

Important as these conceptions are as a point of departure for formulating the relationship between cognition and conation, Gestaltists have not formulated a systematic theory of personality, ego function, and motivation. Nor have they delved into the processes which mediate between motivation and representation. Aside from an incomplete definition of attitude, they have not really shown how attitudes influence the "autochthonous" Gestalten in the phenomenal field. The reason for this may lie in the fact that the Gestalt school has limited its objective to first establishing *formal* principles of representation in the organism-environment relationship. The cognitive aspects of stimulus and response have been developed in terms of organizational structures which represent either the distal or the internal environment. Such an approach may be considered a basic mapping out of what Gibson (1951) has called the "world of literal perception." In contradistinction, the "world of schematic perception" has not been explored. The latter, according to Gibson, includes the significance of objects, their meanings and symbolizations as the person acquires them in his particular cultural milieu. The Gestaltists have not attempted to deal with the acquisition and function of symbols, concepts, ideational contents, and language, nor with the manner in which all these are reflected in the phenomenal field and in motivated behavior. Koffka acknowledges this gap.

> The environment in which we behave is not only the actually and palpably present behavioral environment, but also the environment which we "merely" imagine or think of. This environment is responsible for the greatest achievements of mankind. Without it neither the sciences nor the arts would be possible. Closely connected with this environment is our language and the similar symbolic functions which we have developed. And an ultimate explanation of the problems of thought and imagination will not be possible without a theory of language and the other symbolic functions. But we shall exclude the study of language from our treatise. (1935, p. 422)

Another admitted gap is the social and value-oriented aspect of behavior. Efforts have been made to implement Gestalt theory in this direction. With regard to value, Köhler (1938) has presented a novel approach of combining the phenomenological analysis of "requiredness" with the formal principles of isomorphism. Asch (1952) has extended the Gestalt approach with his cognitive analysis of human interaction. However, in these later efforts, symbolic behavior and its relation to motivation are secondary.

> Problems which arise in this connection may be typified by observations which Bruner (1951) cites. It is only with the greatest difficulty that one can teach or "force" the Chukchee Indians "to carry out anything resembling an adequate sorting of the Holmgren yarns" because of their extremely limited color nomenclature. "Yet these reindeer-herding people can and do apply more than two dozen names to the task of distinguishing the patterns of the reindeer hides" (p. 136). The suggested distinction between literal and schematic perception is here quite apparent.
>
> The influence of cultural and personality factors on cognition, and vice versa, has been the focus of numerous experiments of the so-called "New-Look" type. The study of the "Honi" phenomenon by Wittreich (1952) is a striking illustration because, like many of its kind, it demonstrates that cognitive and motivational factors intertwine but does not explain how. In the Ames (distorted) room, the S commonly sees the room as normal, while a face will shrink or enlarge in size depending upon the window through which it is seen. When faced with his own spouse, however, the S reports the room as distorted rather than the face of his marital partner (Honi phenomenon). The spouse's face appears less changed from its normal size than is that of a stranger. Similar findings were obtained when the marital partner was compared with a stranger walking through the room. Strangers of both sexes were used. Twenty-four of the thirty-two S's were married less than a year and the Honi phenomenon was predominant among them. Familiarity and frequency as an explanation were ruled out, since there was a negative correlation between length of marriage and the measured size of the Honi effect.

The examination of the cognitive aspects of the response points to the problem of silent representation as a behavior determinant. The decisive difference between the cognitive and S-R approach as to the nature of traces may be thus summarized. The associationists conceive them as static and specific, e.g., *habits* which can be combined in various families. Cognition theorists, on the other hand, conceptualize traces as representing relations or systems of relations. In the habit theory, items of knowledge cannot be modified and new problems are solved by reinstating identical or fractional identical past responses. The organization hypothesis permits the restructuring of past responses in a novel way and offers a unitary explanation of transfer and productive thinking. It also introduces a systematic account of how manifest responses occur in a nonconscious context. This context is a *schematic* representation operative as a frame of reference, preverbal generalization, or set, all of which are usually transphenomenal.

In assessing the limitations of the formal-structural approach of the Gestalt school, the question has been posed: If acquired cultural meanings, values, and motives become part of silent organization, how can they be treated from a molar point of view? Taking their lead from Tolman's work, Bruner (1951), and Postman (1951) answer this question by putting forth the construct of "hypothesis formation" as an intervening variable. The individual's predispositions to hypothesize are value-laden and operate selectively in terms of value congruence (or incongruence) of the information provided by stimulus cues. Yet, it seems that the adoption of values is not really analyzed, and motivation is treated only as an experimentally induced or pre-existing variable. Klein and Schlesinger (1949) have criticized this and similar approaches for neglecting the person as a whole. Such methods isolate motivational variables and treat them outside of their dependency on and embeddedness in the individual. There is a manipulation of motives and values in a piecemeal way, implying a constancy hypothesis, which tends to de-emphasize individual differences. Eventually cognitive theory will have to agree on a conception of personality and motivation which reconciles the above difficulties arising from "idiographic" and "nomothetic" emphases.

CONATION AND COGNITION

There is general agreement that man is a wanting creature. There is disagreement as to whether he is a wanting human or a wanting animal. Psychologists who prefer an "animal-centric" conception have linked all motivation to the mechanism of primary and secondary drives. In man no place is given to a genuine interest in the understanding of his surroundings or for cognitive commerce with it. Instead, knowledge is reduced to learning that consists of stimulus-response couplings associated with the satisfaction of primary physiological needs. Any goal beyond the primary needs, as well as the means for its attainment, is derived from additional associations. In principle, motives such as prestige, love, cooperation, and submission are responses which have become instrumentally coupled to stimuli or "cues" during the process of gratifying the basic needs. Such connections between drives and learned behavior are the S-R alternatives to cognitive anticipation and structuring of relevant goals and to the insight into means-end relations of "what leads to what." Although efforts have been made to define reward and punishment nonsubjectively and operationally (Thorndike, 1935), S-R drive theory, in general, has hedonistic roots. As Postman (1947) has pointed out, hedonism is reflected in the current version of the law of effect as "primary reinforcement." This is no longer the classical hedonism in which pleasure is the motive. Postman speaks of "neohedonism," where tension or discomfort is the motive, while pleasure is the result of tension reduction, and it is that relief which leads to learning. Mowrer (1946), in support of the reinforcement principle, equates pleasure, satisfaction, and drive reduction.

Points of dispute between S-R theorists and their opponents have arisen in connection with such a drive theory. The issues may be crystallized by considering the following arguments raised by various cognitive theorists.

1. Are all motives ultimately reducible to a number of specific organic needs? Is the concept of tension reduction an adequate model for human motivation?
2. Is primary or secondary need gratification the only condition for the attainment of information or knowledge?
3. Is the "egocentric conception" of motivation, as implied in neohedonism, veridical? Does it account for both the subjective and objective reference of human goals and emotions as well as for their cognitive contents?

It is obvious that these questions are all intermeshed. Any separation would be artificial. For reasons of practicality, however, we shall have to examine each of them briefly.

PRIMARY NEEDS, TENSION REDUCTION, AND COGNITION

The premise that the biological needs have primacy in man stems from infrahuman and atypical observations. By manipulating deprivation states in animals the existence of independent "viscerogenic" drives or needs is inferred. This is then applied to man, particularly to the human infant. The apparent potency of body needs in infancy is not seen as a symptom of immaturity and lack of integration. Instead, the drive states and release urges in the infant become the paradigm for the motives of the mature adult. Outside of Freudian ideology this model may have its origin in a mistaken notion of Cannon's (1939) theory of homeostasis. From his experiments it became apparent that different tissue needs can disturb the internal equilibrium of the body and lead to specific "hungers." If internal self-regulation itself does not compensate for these deficiencies, then the organism has to restore the homeostatic balance by interaction with the *external* environment.

By making tissue deficiency a *Stimulus* for a motivated *Response*, behaviorists universalized such an S-R relationship for human strivings. This follows the schema: S = Specific organic need state (Tension or Discomfort) ⟶ R = motor activity reducing this state (Tension discharge or Pleasure). It has been said that such a schema applies only to cyclic physiological imbalances and their equalization. Goldstein (1939) as well as Maslow (1943a, 1943b) have argued that human behavior thus actuated is atypical rather than normal. It occurs when the person is governed by acute or chronic deprivation. Such lack or frustration can lead to an abnormal isolation of part processes which then dominate the organism and give the impression of powerful motives. Under normal conditions, there are less severe lacks and the person shows integrated, cognitively directed, behavior. He does not operate on a physiological push-button basis but realizes different purposes and "cognitive needs" (Maslow, 1948a, 1948b). The transient nature of manifest drives stemming from homeostatic imbalances is for Goldstein a cogent argument against elevating them to the status of primary and permanent motives. The need for salt, for example, cannot be considered a true motive, unless there is craving for it because of abnormal deficit. Activities connected with homeostatic maintenance are merely instrumental and subservient to the all-encompassing drive of the human organism — the optimal realization of his capacities. The motives corresponding to these capacities have a new quality: they can have cognitive content and object reference. Thus fulfillment of most organic needs enables the human organism to actualize his characteristic potentialities. Such need fulfillment can even become an avenue of self-expression. Sex gratification, for example, becomes part of realizing a capability for the emotion of love. Both Rogers (1951) and Fromm (1947) have given the actualization of potentialities a prominent place.

Cannon himself was the first to realize the restricted nature of the deficiency approach to human motives and emphasized the liberating function of homeostasis *from* the concern with the details of bare existence *for* the activity of the higher levels of the nervous system. "The organism is released from the limitations imposed by internal and external disturbances . . . liberated for its more complicated and socially more important tasks . . . free to enjoy beautiful things, to explore and understand the world . . . to develop new ideas and interests" (1939, pp. 302, 303, 322 ff.).

From such a point of view it would follow that when bodily needs are satisfied, human motivation cannot come to a standstill. There is much truth in the adage: the higher the standards of living in a given society, "the more people eat in order to work rather than work in order to eat." S-R adherents have explained such phenomena by grafting upon the primary need reduction secondary rewards of coincidental associations, any one of which can become a substitute incentive for primary needs.

A further limiting feature of this derivation is that it conceptualizes primary or secondary need gratification only in terms of antecedents and consequences without considering a phenomenal representation of the need states and the end states (cf. Young, 1952). From a phenomenological point of view, some writers have criticized reinforcement theorists, and Freud too, for treating motivation primarily as a negative force, i.e., defining it as a tension-reducing vector. On this point there is convergence among authors such as Allport (1946a, 1947), Goldstein (1939, 1940), Schilder (1942), Hartmann (1951). Each, in his own theoretical version, has been prompted to give a prominent place to positive aspects of both the *setting* of goals and the consummatory phase of their attainment. In the first place, the individual is not always engaged in avoiding (reducing) tension. He is just as often creating tension in terms of tasks, interests and long range objectives. Second, the pleasure component in achievement is not altogether a relief from tension nor is it a mere end state of a former disequilibrium. More than this, goal attainment is usually a *cognitive* enrichment which leads to a changed, differentiated state of the person (Allport, 1946a; Piaget, 1951). This new state of the person allows for a widened horizon in which new goals, previously not realizable, can become envisaged. A comparison between the drive reduction tenets of the Freudian or the S-R type and the cognitive alternatives cited here may be illuminating. In the former, motives are conceptualized as reactions to disturbances resulting in a "state of rest." In the latter, motives are not discharges of tension. The person spontaneously creates objects of endeavor and activates purposes mediated by cognition. Furthermore, the person does not remain static but progressively changes with each goal attainment.

The productive possibilities in treating motivation in this framework may be illustrated by the thought-provoking formulations of Schilder (1942). He asserts that living organisms, particularly the human, have the drive not to remain static in "completeness," but to go "from completeness to incompleteness to new completeness." Applied to the sex drive, for example, "the psychic expression of sexuality is incompleteness and the necessity for contact with another individual. Sex is, therefore, one specific expression of the 'social' characteristics of life" (p. 280). What is completeness in prospect of a goal becomes incompleteness after reaching it, because the person has now become more differentiated. One reason for the change in the person following attainments is suggested by Maslow (1948b), in agreement with the Gestaltists. There is a nonarbitrary relationship between the person and the satisfying object, in that it has intrinsic appropriateness. This implies that some contents of the environmental goal interact with the behaver in an objectively fitting way. He must acquire an internal representation of both the means and the attained end. In turn, this reorganizes the person and his life space. Both Woodworth and Hilgard have called attention to a motivational factor in perception which is virtually antithetical to tension reduction. Woodworth's thesis is:

> Perception is always driven by a direct, inherent motive which might be called the will to perceive. Whatever ulterior motives may be present from time to time, this direct perceptual motive is always present in any use of the senses. It is impossible to look without trying to see or to listen without trying to hear. To see, to hear — to see clearly, to hear distinctly — to make out what it is one is seeing or hearing — moment by moment, such concrete immediate motives dominate the life of relation with the environment. . . . When the goal of such a search is attained, strong reinforcement is revealed by the observer's cry of satisfaction and later by his excellent retention of the discovered figure." (1947, p. 123)

Correspondingly Hilgard (1951) speaks of two goals of perception — one, to attain stability in the world about us, and two, to achieve clarity in what we perceive. Granted that these perceptual goals are also attended by satisfaction and may be embedded in wider ends of the person, they are quite remote from any association with drive reduction.

The emphasis on the cognitive facets of motivation is thus clearly in opposition to primary and secondary drive theories. One may epitomize the various arguments in this reasoning: the more the human organism is freed from the immediate urgencies of tissue hungers, the more potent do the inherently human and social potentials in motivation become. Thus, Maslow (1948a) has proposed a hierarchy of "higher" and "lower" needs in which, for example, the "living at a higher need level" leads to even greater biological efficiency and mental health. The fulfillment of the higher needs is closer to the actualization of one's capacities than is the gratification of lower needs.

In contrasting drive reduction with capacity fulfillment, we have mainly cited the following points: the individual strives for cognized objects and goals, not for satisfaction itself; having attained a specific goal, he sets new ones for himself. Nothing has as yet been mentioned about motivations which transcend the phenomenal field or goals which are not represented in it — nonconscious motives. There is actually no provision for a distinction between conscious representation and unconscious motivation in drive reduction theories of the S-R type. As for the Freudian drive-reduction theory, unconscious long-range persistent strivings in the person are converted into repetitions of childhood frustrations. Orientation towards the future becomes re-enactment of the past. Granted that clinically such motivations are often reported, it should not detract from the equally legitimate occurrence of future oriented behavior. Allport (1947) has given this problem recognition with his concept of "intention" as a value-directed activity. An intention is a guiding idea which may be in or outside of awareness. "In my use of the term I do not confine intention to reportable purpose. Sometimes the essential direction of an intention is understood well enough by the subject, sometimes not. If the term is taken to mean *both* the understood and the nonunderstood direction of an act . . . it can serve as a proper designation for 'ultimate' motives, not merely for proximate or rationalized motives" (p. 187n). Allport distinguishes intention from "anticipatory goal reaction" as well as from "expectancy" because:

> Infant and rodent have immediate goals and indulge in anticipatory goal reactions but have no directive schemata . . . Men often have values without having any specific goal in mind. They may have a consistent direction of striving, but their goals are either trans[illegible] or else undefinable. All of a rat's, but only digmall bit of human, behavior can be characterized in terms of concrete goals whose attainment will de-tension specific drives. For the most part, the course of man's behavior runs according to certain schemata, or in prolonged channels. Only now and then are these channels marked by lights or buoys that represent specific goals . . . We may know a person's expectancies and even his past rewards, and yet be singularly unable to predict or control his future behavior, unless at the same time we know also his basic intentions, which are by no means a stencilled copy of his previous expectancies and rewards. (p. 184 ff.)

This approach embodies the positive direction of motivation and at the same time leaves room for its internal representation, which may or may not be in direct experience. Allport's conception of *directive schemata* encompasses these factors. It is noteworthy that his formulation connects a motivational vector with silent organization.

Essentially "intention" means "*what* the person is trying to do" whether in or out of awareness. In certain respects we have a supplement to this in Klein's (1952) analysis of *how* the person does what he is trying to do. Extending his former work on personal styles of the perceiver, Klein experimentally explored the relationship between need tension and cognition. He focused on the manner in which the ego handles needs, and he used thirst conditions as a paradigm for need arousal and ego processes: S's were given, among other tasks, a number of perceptual problems, when thirsty and when sated. As one function of the ego structure the experimenter distinguished "cognitive attitudes" which have regulative and directional properties in the control and expression of needs. Briefly, his thesis is that people are organized differently to deal with needs in that they have individual styles of *cognitively* channeling them.

Klein does not equate cognitive controls with defense, which for him is not the prototype of all control. Defenses are modes of resolving conflicts among needs, whereas cognitive attitudes might even be called "conflict-free styles of regulation." This conclusion follows from his experimental findings. The cognitive styles tapped in his S's also influenced their perceptual structuring of situations where quite "neutral" stimuli were involved. Furthermore, the selective effect of cognitive attitudes on the S's responses was not linked to "satisfaction-giving objects." Cognitive attitudes are adaptive in the sense that they determine the manner in which needs are expressed in conformance with the task, i.e., reality, but they are not necessarily tension reducing. As a matter of fact, one group of S's actually was more accurate in its performance of tachistoscopic recognition under need-tension than when sated. While thirst "lamed the overcontrolled" S's, it seemed to "wake up" the opposite type of S's, whose over-all errors decreased in the various tasks. Instead of dealing with specific contents, as needs or defenses do, the cognitive attitudes express themselves in formal qualities of behavior, e.g., "the particular *way* that the stimulus field is organized." Cognitive controls are indicated by such processes as "*schematizing* or conceptualizing behavior, *mapping, absolute and differential thresholds, anticipation, attention, concentration*, etc." In the light of Klein's formulations, certain of the experiments on the influence of needs upon perception may take on a new look. It is his contention that by treating need intensity as an independent variable and a perceptual response as a dependent variable, the ego structure is overlooked. The contradictory reports about the direction of cognition under the influence of need states is taken by Klein as support for his claim that cognitive attitudes are "intervening variables." And these are necessary for an explanation of the individual differences and consistencies in such experiments.

It is true that the "new look" experiments have revived interest in the interrelatedness of motivation and cognition. However, they have sometimes oversimplified that relationship. Cognition was dealt with as if it were object perception along an external stimulus dimension correlated to a motive along an internal stimulus dimension. The experimentally induced distortions are then usually obtained under reduced or equivocal stimulus presentations. Since the deviations are rather small and inconsistent under these conditions, generalizations about a one-to-one relation between need states and sensory perception may be premature. Even such discrepancies as found in the coin-size experiments do not really shatter the veridicality of ordinary object perception. It has been further questioned whether the "vise-like grip" of needs significantly distorts our percepts (Pratt, 1950; Luchins, 1951; Gibson, 1951). The use of equivocal stimulus conditions for tapping the personological aspects of perception seems to have been more profitable in settings such as the following: (1) When the consistency of intraorganismic processes of the person's functioning is investigated rather than a direct motive-percept correlation (Werner and Wapner, 1952; Klein, 1952; Witkin, 1953). (2) When the values that are pertinent to a person (including self-regard) are studied in relation to stimulus material which belongs to the realm of learned meanings, symbols, and cognitive contents (Postman, Bruner, and McGinnies, 1948; Postman and Schneider, 1951).

REINFORCEMENT AND COGNITION

The doubts cast on primary drive-and-tension-reduction as a suitable model for motivation necessarily implicate the reinforcement principle. The well-known controversy on latent learning is a further expression of this state of affairs. The question is whether reinforcement is the sole condition for attaining knowledge. Both the defenders and critics of latent learning have tried to decide the issue, chiefly through animal experimentation. Typically, rats first explore a maze without getting food reward and are then tested for learning this maze with reward introduced. Since they are superior to a control group which had not previously explored that maze, advocates of learning sans reward see this as a support for their theory. Since Tolman's students introduced this method as a demonstration of cognitive mapping without primary reinforcement, numerous variations of the experimental design have been carried out, with varied results.

In evaluating the experimental evidence for latent learning, Hilgard (1948) concludes that it is "severely critical of the reinforcement principle." Thistlethwaite reviewed the research in this area and accepted a number of experiments as "constituting exceptions to the alleged indispensability of reinforcement in the learning process" (1951, p. 125). The issue has, however, not been closed, as the continued debate over *concealed* variables of noncognitive nature attests. Thus most investigators agree on the superiority of the *performance* by "latent learners," but they disagree as to the drive and response variables involved in the *learning process itself.*

Tolman's latest stand on the matter is that latent learning is ushered in by a motive, namely, a "curiosity" or "placing need." New "field relationships" are learned through a motivation which is a "relatively 'disinterested' one provided by a general exploratory need" (1951, p. 304). The gratification of such an exploratory motive during latent learning seems to parallel the satisfaction which Woodworth and Hilgard ascribe to the attainment of clarity and stability in perception. For Tolman, such disinterested cognitive mapping can become a vehicle for a specifically acute need aroused at a later time.

Notwithstanding the admirable ingenuity which has gone into the designing of crucial experiments in this area, psychologists have not sought to resolve the entire issue of reinforcement on phyletically higher levels. This would entail the difficulty, if not impracticality, of testing for reinforcement in man by primary need deprivation in an experimentally "pure" way. After all, inexperienced human infants have never been deprived in learning experiments, as have been rats. Therefore the general application of primary reinforcement to the human level is not operational but speculative in method. In this regard Harlow's (1953) experimental critique of reinforcement deserves attention. If primary drive reduction is not a necessary factor in the learning of monkeys, then reinforcement becomes questionable as the sole principle of learning in man. Harlow's data indicate: rhesus monkeys learn efficiently even though they were fed previously and munched stored food during both right and wrong responses; they solved complex mechanical problems spontaneously with no other incentive than the presence of the puzzle; investigatory motives can be more potent than hunger tension. As an alternative to primary drive-reduction theory, Harlow offers the hypothesis that responsiveness to external stimulus patterns is decisive for motivated learning and increases in importance with the complexity of organisms. The key to human learning for Harlow is not the conditioned response. He asks "can anyone seriously believe that the insatiable curiosity-investigatory motivation of the child is a second-order or derived drive conditioned upon hunger or sex or any other internal drive?" (p. 29).

SECONDARY REINFORCEMENT AND COGNITION

The preceding discussion has alluded to the principle of secondary reinforcement. It has been applied to processes of socialization and the acquisition of attitudes. The radical difference between cognitive and S-R approaches to "conditioning" of likes and dislikes may be illustrated by three typical instances.

1. Among the classical studies is M. C. Jones' (1924) experiment on the "de-conditioning" of the boy, Peter's, morbid fear of rabbits. The schema of conditioning the food response to the rabbit is still being diagrammed in many textbooks as a paradigm of secondary reward learning and stimulus-response substitution. In the "direct conditioning" the experimenter brought the rabbit as close as she could without interfering with Peter's eating. After several sessions, Peter was apparently cured, explaining, "I like the rabbit." Concerning the substitution theory, Tolman has satirically wondered why Peter does not now eat the rabbit. Hence a *joy*-over-food response has been "substituted" for the food response itself. Even this hypothesis raises questions: Has the original response remained the same and an additional response merely become attached to it? Will Peter henceforth make one response to "dangerous rabbit" plus a more potent joy-over-food response? Along the lines suggested by Tolman (1932) it seems possible, however, that the original stimulus changed from dangerous rabbit to harmless rabbit. If this be true, then the role of the eating pleasure was not that of a pacemaker but of a peacemaker in a cognitive restructuring. The joy-over-food was a pleasant distraction during the rabbit presentations, which provided an opportunity to *discover* that the rabbit was harmless.

This discussion, however, is only based upon the idealized version of the Jones' experiment. Her careful protocols reveal important facts frequently overlooked. The entire training lasted several months. The first seven sessions took place in the presence of three other children who "were selected carefully because of their entirely fearless attitude towards the rabbit" (p. 310). "Peter progressed from a great fear of the rabbit to a tranquil indifference and even a voluntary pat on the rabbit's back" (p. 312). After two months' interruption, his fear re-appeared when a dog frightened both Peter and his nurse. Now the method of "direct conditioning" was introduced and "occasionally also, other children were brought in to help with the 'unconditioning' " (p. 313). Jones noted most improvement on four occasions. In these, either other children or an assistant whom Peter called "Papa" were present. The first improvement was recorded after Peter had followed and watched a child playing with the rabbit, the second when the rabbit was given to another child and Peter wanted it on his lap. These data contraindicate a mere substitution of stimuli or responses. The variable, joy-over-food response was *contaminated* by the variable of perceiving fearless and unharmed children playing with the "dangerous" rabbit. The introduction of such social interaction makes it doubly necessary to consider cognitive factors in Peter's loss of fear. This, of course, could be answered by stipulating that he learned through "imitation."

2. In studying imitation in children, Miller and Dollard (1941) on the one hand and Asch with coworkers (1952) on the other have employed similar experimental designs. Both tested how children follow a leader who found candy in one of two boxes. The only difference was that in the experiments reported by Asch the "followers" were interrogated after each trial ("How did you know which box to go to?"). There are important contrasts in the respective conclusions. For Miller and Dollard the leading child is the cue for the followers' response of finding the rewarding candy in one of the two boxes. By prearranged design the follower always found the candy either in the box identical to or opposite to the leader's choice. In the identical choice the leader is a positive cue for reward; in the "opposite" choice the leader is a negative cue. No knowledge of the why or wherefore is assumed in this learning, and since the prearrangement was not known to the followers, conditioning by association without understanding appears to be operative. This principle is then universalized for all imitation in leader-following (1941, p. 97). Asch, however, using the verbal protocols, found that the followers think. They make "assumptions that serve as hypotheses, which they formulate, test, revise, reject, or confirm" (p. 394); they attempt to reach an understanding of the hidden system. When the candy in the box was replaced by a rubber ball which had to be returned after each trial, the same, if not more serious, efforts on the part of the children were observed. They apparently were more interested in finding a rubber ball not-for-keeps than a candy-for-eats.

An intriguing question arises out of Miller and Dollard's data. Why did the children average only one-fourth of the errors when they learned to go to the "opposite" as compared to the "identical" box? And why did 77% of *all* children go to the box opposite the leader's choice in the *first* trial in *all* cases? Could we not say they reasoned by generalizing from past life experiences: "One candy has already been taken away by the child ahead of me from *this* box; the candy *for me* will be in the *other* box?"

In comparing the S-R with the cognitive interpretation of imitation, Asch criticizes the concept of "cue." It is insufficient to call a perceived action a cue. One has to state the perceptual conditions under which certain parts of the action and not others are singled out as relevant. This reintroduces the factor of social perception and phenomenal causality. It also points to the reasoning necessary for learning from another person's experiences. This applies to the case of Peter as well as to the above case, where learning from the leader's reward (which is not one's own reward) has to occur.

3. In the area of attitude formation, cognition advocates have increasingly stressed the role of central and representational processes. Both McLeod and Asch, for example, concur that the ideational content of an attitude has to be included in psychological analysis. Asch speaks of attitudes as "an organization of experiences and data with reference to an object" (1952, p. 580). McLeod asserts that an attitude has both subjective and objective components. "If we are to characterize adequately an attitude towards 'religion,' we must specify both what is objectively there as symbolized by the term 'religion' and what is subjectively felt in relation to that objective structure" (1947, pp. 202-209). Moreover, attitudes are conceived as value directions which are organizationally related to each other in a systematic way. Following the reasoning of G. W. Allport (1935), critical reservations have been made about treating attitudes as particularized responses to specific stimuli. Thus McLeod cautions against the atomistic tendency to hypothesize a separately measurable attitude in correspondence to each *name* which designates a social object, such as race, class, nation, etc.

The aforementioned comments together with the recent formulations of Adams (1953) and Heider (1953) converge on two characteristics of attitude. Cognitively, it organizes and orders certain events into a unified view, a representational structure. The emotional or subjective nature of attitude consists in its functional significance for the person in supporting a stable relationship between him and his social field. In this role it is interdependent with other attitudes, forming a system. It is for this reason that the attempt to influence attitudes by dealing with any one of them singly often meets with such strong resistance. A change in one part of the system "threatens" the entire system, so that the resistance of a particular attitude derives its strength actually from the whole structure of which it is a part.

In contrast, reinforcement theorists have tried to deal with attitudes as specific responses which do not differ essentially from any other S-R connections that are either closely or indirectly associated with reward. Miller and Dollard (1941) as well as Doob (1947) are outspoken exponents of this trend. Attitude becomes a learned cue which is tension producing and evokes responses as anticipatory goal reactions which lead to tension reduction or reward. This view is reflected in such statements as "Most people learn to love those persons, objects, places, and conditions which are associated with less hunger, less thirst, less pain and fatigue or more money and prestige" (May, 1944, p. 51). We are dealing here with an S-R chain of responses in which each response evokes further acts that bring about a reinforcing state of affairs. It is true that the origin of many motives has been located in the learning of subgoal activities while pursuing primary need gratification. This does not dispose, however, of the role of cognitive processes and changes in such learning. For example, that a love attachment to the mother accrues in the child from the discomfort relief by the mother's nursing is a common tenet advanced in this context. Even if true, it is difficult to accept the assumption that this love attachment remains unmodified by cognitions in the maturing individual. Otherwise the reversal from the infantile attitude of loving mother because-she-nurtures-me to the adult attitude of nurturing mother because-she-needs-me is hard to explain, especially when the average adult is not wholly uncritical of the average mother. The theory also conflicts with the adoption of attitudes which have not been rewarded in the past and do not offer reward in the future, as in the case of conscientious objectors and other nonconformists.

Upon closer inspection, the concept of secondary reinforcement appears to be a rather weak, if not elusive, link. There is too wide a gap from the presumed S-R bonds during primary drive reduction to the more complex, remote S-R chains which involve the learning of subgoals and their development into self-sustained social motives.

Whereas G. W. Allport tried to solve the problem of an increasing independence of former subgoal motives by his concept of "functional autonomy," Hull tried to replace both cognition and motivation by contiguity in association. Hilgard's judgment is:

> The integration of behavior choices remote from a tension-reducing goal object is usually accounted for on the basis of some sort of internalized stimulus-response sequences, supported by secondary reinforcements. Hull . . . finds it necessary to assert that secondary reinforcing agents need not bring about tension reduction in order to be reinforcing. This is almost a complete concession to cognitive control, and weakens greatly the centrality of the reinforcement principle. In describing the essential similarities between secondary reinforcement theory and cognitive theory it is clearer to state that Hull reformulates the cognitive theory than to assert that the cognitive theory is just another way of talking about secondary reinforcement. (1948, p. 345)

MOTIVES, EGO, AND COGNITION

It may be said that the doctrine of drive reduction embraces hedonism in the form of tension release, while the same doctrine rarely postulates an ego that can "enjoy" its need reductions. In turn, if an ego or self is posited, personality is viewed as being governed by self-interest qua tension release. For example, in orthodox Freudian thinking, cognition becomes essentially subservient to self-centered motives, and genuine motives cannot be induced by cognitions.

The state of affairs may be illustrated by the controversy between Allport (1946a) and Mowrer (1946) on the ego and the law of effect. Mowrer recognizes the necessity for an ego by postulating ego processes in reward learning, but would like to translate ego processes into "emotional arousal." Emotional reactions as well as " 'interests' are secondary drives and . . . as such they can both motivate and reinforce behavior, in much the same fashion that primary drives do" (p. 327). Thus, interests are "covert" emotional responses and are learned in essentially the same way as other responses. In contrast, Allport considers the ego a fundamental construct for motivation. Motivation in learning consists of interests which belong to the ego structure. Reward becomes just one of many factors for perceiving the consequences of an act and for applying that knowledge in the future when a similar act or situation occurs. Satisfaction or dissatisfaction may at best serve as indicators of success. And success is always evaluated according to its relevance for the interests that the ego pursues. In attacking the universality of the law of effect and defending an antihedonistic conception of learning through interests, Allport presents these arguments: there is no evidence for self-administered "covert" reward (i.e., "internal O. K. reactions"). The inference of such reward is based on a circular reasoning — because a person persists in his activity it is concluded that he has been rewarded. "Interests often persist in the face of continuous dissatisfaction or failure" (p. 346). The tenor of Allport's reasoning suggests that the concept of interest is related to that of intention; both constructs share the openness of a general directional system, as distinct from behavior guided by a specific goal channel. Interests seem to embody two components: a cognitive relationship of the person to the environment, and a subjective relationship to his ego values. The interests are more central to the person and therefore lead to his active participation in pertinent activities and contents. Interest satisfaction in learning is not predominantly "emotional" in the sense of tension reduction. "Learning proceeds because it is relevant to an interest-system: it adds to knowledge, it differentiates items within the system, it broadens the range of equivalent stimuli" (p. 346). It is therefore the person who learns, who is rewarded, and not a single response or concatenation of responses; and it is the person who utilizes the reward as a cue for his progress in the learning activity.

Once it is accepted that an ego construct is necessary to account for both cognitive and motivational directions, then the relationship between motivation and cognition takes on a new aspect. One may now deal with this relationship "from above" rather than "from below." Instead of postulating need pushes which drive the ego, one can conceive of a less direct, holistically mediated determination of behavior. Motivation may be mediated via reference to the ego and via evaluation by the ego. It follows that such mediation has to take place centrally and involves internal representation. One of the outstanding contributions in this direction has been made by Rogers (1951). The point of departure for this new emphasis may be typified in the following statement: "The only channels by which needs may be satisfied are those which are consistent with the organized concept of self" (p. 508).

Basic to his theory is the cognitive relationship between the organism and the environment. Rogers considers an S-R explanation of behavior inadequate. Instead, he asserts that the person reacts as an organized whole within his phenomenal field, which in turn is a perceptual "map" of the outside world and himself in it. He posits figure-ground organization for this field. There exists also a wider part of experiences, connected with the internal and external environment, which is outside of awareness and comparable to the behavioral environment of the Gestaltists. The central theoretical construct is the self-concept of the individual. Although Rogers elaborates a more precise definition of the self, in a general way the self may be thought of as all the things the person consciously thinks he is. The term "self" therefore is not synonymous with organism, but is used to denote "the awareness of being, of functioning" (p. 498).

The structure of the self is formed "as a result of the interaction with the environment, particularly as a result of evaluational interaction with others"; it is "an organized, fluid, but consistent conceptual pattern of perceptions, of characteristics, and relationships of the 'I' or the 'me,' together with values attached to these concepts" (p. 498). There are two ways in which this process occurs. One way is through "direct organismic valuing," i.e., through the experiences which the individual absorbs from his own sensory and visceral evidence. This is raised to the level of *symbolization* by attributing the experience to himself. For example, there is the dawning awareness of the infant, though nonverbal, of "I like," "This is cold and I dislike it." Such direct experiences become part of the concept of self. A second way is the introjection of experiences or values of others *as if* they were one's own. This occurs in an effort to keep new experiences consistent with the already existing self-concept, and is accomplished by distorting the meaning or symbolization of the experience.

> For example: A child feels loved by his parents because of all the direct evidence which comes from being cuddled, talked to, etc. But then the parents reprimand the child when he hits his baby brother and he is made to feel, "You are not loved when you behave this way." In order to preserve the threatened concept of the self as lovable, the child denies to awareness the satisfaction that is experienced in the aggressive act. In place of a more accurate symbolization of the anger expression as satisfying, the experience becomes distorted into the symbolization "bad," "I perceive this behavior as unsatisfying to me," "I do not like to hit baby brother." The concept of the self now includes "I am lovable," and the distorted symbolizations "I like baby brother," "anger is bad."

Akin to the Gestalt concept of fittingness, the principle of congruence or consistency with the self-concept is therefore a controlling factor of behavior. As an example, Rogers notes the man who, possessing a self-concept of honesty, cannot seek achievement through means which seem to him dishonest. The same tendency can lead to distorted symbolization, as in the case of the hitherto brave pilot, who, overcome by fear, cannot simply flee, as this would be inconsistent with his self-portrait of a fearless person. Instead he falls ill and has to be excused from the mission. In cases where the person realizes that a behavior is inconsistent with the self-concept, as sometimes happens when the behavior is "forced" by the pressure of strong needs, then it is felt as not "owned" by the individual — "I really wasn't responsible for what I was doing" (p. 509).

The representational map of one's self and of the environment is best understood from the internal frame of reference of the individual himself. He is considered the center of his experienced, "private" world or psychological reality. This reality, which corresponds to the behavioral environment, can be responded to in one of three ways: " (a) symbolized, perceived, and organized into some relationship to the self, (b) ignored because there is no perceived relationship to the self structure, (c) denied symbolization or given a distorted symbolization because the experience is inconsistent with the structure of the self" (p. 503). In his concept of psychological reality as dependent on the frame of reference of the individual, Rogers does not go as far as Snygg and Combs (1949), who dispense with reference to an organism or to neural processes in relation to external stimuli. The individual can test out perceptions which arise from the interplay between external reality and neural processes (p. 486). The total psychological reality therefore consists, for Rogers, of a combination of testable and untestable "hypotheses" or cognitions.

A certain limitation in Rogers' thinking seems to stem from his tendency to make the individual the center of a "private world." One can readily see that in a therapeutic orientation, understanding the client as being the center of his world is one of the essential vehicles for clinical action. To form a general theory of behavioral reality around this "center" may, however, invite an unintended, mistaken emphasis on ego-centered motivation.

The notion that not all motives or emotions have ego reference has found increased backing (Koffka, 1935; Köhler, 1938; Allport, 1943, 1945; MacLeod, 1947, 1949; Asch, 1952). A typical statement is one by MacLeod. "If we try to envisage all motivation as self-generated we are doing violence to some of the more obvious facts of observation. The vectors may emerge from the object just as validly as from the self. The self may be remodeled to meet the demands of the object just as truly as the object may be shaped to fit the needs of the self" (1949, p. 65). As an experimental illustration of what the above writers have in mind, we may take H. B. Lewis' (1944) findings that "my" tensions towards task completions, when interrupted, can be resolved if a teammate completes them. The suggested distinction between task and ego orientation opens the possibility for dealing with group-centered and other-centered behavior as being intrinsically motivated. No *ad hoc* or *post hoc* hypothesized mechanisms of ego expansion, identification, etc., would be necessary to account for behavior that transcends the orbit of ego-centered motivation.

Task orientation as a nonego-referred direction of motivation seems to demonstrate the "openness" of the ego to the world, a perceptual and cognitive grasp of environmental contents. From this point of view, McLeod's proposal warrants attention: "what we call motives are systems polarized in the dimension of subjectivity-objectivity" (1949, p. 65). The consequences of such a definition would be that, on either pole, the subjective or the objective, emotion and cognition are inseparable. This raises the question whether the traditional separation between conation and cognition, or the customary interaction theory, gives a realistic picture of behavior. Most cognitive theorists have moved towards a more unified treatment of these functions. This was already implied in the rationale of such concepts as interests, intentions, cognitive attitudes, self-concept, etc. Representational and ideational characteristics have been attributed to conations just as much as motivational direction and emotional quality have been claimed for cognition. Thus, writers on this subject nearly all see an inherent connection between striving, feeling, and knowing. (Take, for example, Tolman's concept of "belief-value-matrix," 1951.) Yet the problem of how to conceptualize this unity and how to demonstrate it is still unresolved.

The most promising avenue of solution seems to come from a phenomenological analysis of emotions which is followed up by an empirical search for the underlying intraorganismic processes. Emotions may well turn out to be the unique expressions of the unity of cognitive and conative functioning. The positive and nonarbitrary nature of emotions, as against their disorganizing and disrupting influence, has been emphasized in recent writings (Leeper, 1948; Goldstein, 1951). The most poignant formulation is given by Angyal, who defines "the feeling tone of emotions as the experience of the state and of the situation of the person under the aspect of value" (1941, p. 71). It appears that any experience or act of valuing presupposes a cognitive-emotional-motivational *matrix.* A stimulating exploration of this matrix can be found in Asch's treatment of emotions (1952, p. 79 ff.; 109 ff.). To have an emotion is to perceive agents acting upon us or others, i.e., to apprehend causal relations. Emotions are, as a rule, directed towards their perceived source — persons or objects. We are angry at or because of, sorry about, happy over, grieving for, and so on. We seem to "know" with what emotion to respond to a situation, i.e., emotions usually have an appropriateness about them. Emotions are the way we react directly to the relations between given conditions and ourselves. They can represent to us the fate of our goals; they are immediate cognitions of the state of affairs of our motives. Once these propositions are accepted, *our* perception of emotions in *others* can represent to us their intentions and motives in an immediate, understood way. It follows that we can emotionally participate in other people's fate. Asch's contention that emotions have a cognitive basis becomes crucial on the level where the individual has developed the idea of a self. This leads to the generation of new and more complex emotions, all hinging on the conceived relation between the idea of myself and others. Such human emotions as pride, envy, gratitude, shame, jealousy, indignation, remorse, embarrassment can now be actualized. Asch makes the interesting point that all these emotions presuppose past experiences and a knowledge of the contents of situations. However, the emotions themselves cannot really be learned or taught. The person has to "generate" the emotion himself on the basis of what he has learned as well as understood about the relevant factors which call for gratitude, remorse, and so on. "There are no 'unconditioned stimuli' to remorse" as there perhaps are to fear (p. 81). Hence there can be no conditioning to remorse. Remorse, like gratitude, goes hand in hand with cognition. Similar reasoning has been applied to the emotions of shame and embarrassment by Buytendijk (1950) and Nuttin (1950) independently.

Other phenomenologically oriented thinkers have emphasized the object reference of emotions, as in the case of "physiognomic" experiences: the feeling of the mood of a landscape, its friendliness or unfriendliness. These and similar object-oriented emotions may also lead to motivation. For example, the individual may feel prompted to take action because he is "moved" by the experience of pity, sympathy, indignation, injustice, etc., in *reference* to a person or a cause (MacLeod, 1947, 1951; Köhler, 1938).

In principle, then, behavior may be conceptualized as being embedded in a cognitive-emotional-motivational matrix in which no true separation is possible. No matter how we slice behavior, the ingredients of motivation-emotion-cognition are present in one order or another. As Adams succinctly expressed it, "For there is a conative or dynamic component in our most disinterested, scientific, and objective cognition, or we would not make the observation. There is likewise a cognitive component in our blindest lust or rage, or their expressions would not have even the minimum of direction that they seem invariably to have" (1953, p. 55). This intimate relationship could not escape the profundity of William James:

> The contrast is really between two *aspects*, in which all mental facts without exception may be taken; their structural aspect, as being subjective, and their functional aspect, as being cognitions. In the former aspect, the highest as well as the lowest is a feeling . . . In the latter aspect, the lowest mental fact as well as the highest may grasp some bit of truth as its content . . . From the cognitive point of view, all mental facts are intellections. From the subjective point of view all are feelings . . . And then we see that the current opposition of Feeling to Knowledge is quite a false issue. If every feeling is at the same time a bit of knowledge, we ought no longer to talk of mental states differing by having more or less of the cognitive quality; they only differ in knowing more or less, in having much fact or little fact for their object. The feeling of a broad scheme of relations is a feeling that knows much; the feeling of a simple quality is a feeling that knows little. But the knowing itself, whether of much or of little, has the same essence, and is as good knowing in the one case as in the other. (James, 1890, Vol. 1, pp. 478-479)

In dealing with the cognitive-emotional-motivational matrix, a kind of leitmotif emerges. Even when the unity of this matrix is recognized it has a polarity of direction. In James' terms, there are the structural (subjective) and functional (cognitive) aspects; in the terms of the Gestaltists and McLeod there is a subjective and an objective pole; in the terms here used there is an ego reference and an objective reference.

There has been a tendency to deal with the subjective pole in terms of the individual *uniqueness* and *privacy* of the person's experience. Conversely, the objective pole has often been equated with the *public* mode of experience, for example, in terms of "intersubjective protocols" as a basis for science. In a sense this difference is also reflected in the methodological debate between the idiographic and nomothetic schools of thought in psychology. A developmental approach which examines how the difference between the subjective and objective poles emerges may throw light on this problem. It may lead to a better understanding of how, for example, the awareness of subjectivity and self-reference arises together with the segregation of the self from its unreflective embeddedness in a social and behavioral surrounding.

COGNITIVE DEVELOPMENT AND GENETIC REDUCTIONISM

Genetic reductionism takes two forms. The first is phylogenetic; the second is ontogenetic. In phylogenetic reductionism, evolutionarily higher stages, e.g., complex human behavior, are treated as if they did not differ fundamentally from lower stages. The difference is quantitative: the higher levels are more complex only in the sense of containing more "elements." No qualitative change is attributed to this increment of complexity. The tendency therefore is to consider human cognition and intelligence in terms of an increased repertoire in the number of associative connections, and not as a new function in the kind or direction of thinking.

In ontogenetic reductionism, similar additive mechanisms are posited for human behavior. Adult behavior is said to be derived from early acquired reactions. The basic response dispositions are inculcated during childhood. The mature personality becomes a "stencilled copy" of previously conditioned motives which have become attached to additional environmental stimuli. In principle, the development of the individual does not create new motives during various genetic stages. Ideational processes — when admitted — are relegated to past learning rather than to the developmental emergence of capacities for reasoning or to new modes of learning. The two types of reductionism are usually complementary, and the second form, the ontogenetic, is often derived from the first, the phylogenetic. This relationship is manifest in those approaches to cognitive phenomena which have the intent of reducing or relating them to S-R laws.

UNIFORM MODES OF LEARNING

It is probably no overstatement to say that behavioristic learning theories have more or less taken it for granted that the laws of learning are the same for different phylogenetic levels and that the same laws hold for different levels of human development. "The extreme form of generalization from comparative studies is that there are no differences, except quantitative ones, between the learning of lower animals and primates, including man" (Hilgard, 1948, p. 329). It has been pointed out by Schneirla that this reasoning leads to the fiction of understanding the various learning phenomena as based on "homogeneous agencies" rather than on "processes of very different nature still requiring explanation" (1949, p. 248). Moreover, human learning is typically not studied in its own form but is extrapolated from lower forms — particularly rat learning. The importance of differences in *capacity* and *kind* of learning has been pointed out by Werner (1940), Schneirla (1949), Harlow (1953), and Tolman (1949). Schneirla has criticized Hull's S-R doctrine for its neglect of such differences and also for its reliance on the conditioned response as the prototype of learning (1949, p. 267). Serious doubts about the universality of the rat model and the premise of uniform learning mechanisms in different species have also been voiced by Birch and Bitterman (1951). Hull (1952), in a posthumous publication, states that most of the postulates are based on the behavior of lower animals, particularly the rat. "Humans have the added capacity of speech, symbolic behavior, with its accompanying advantages to the higher mental processes. Whether this introduces any primary behavioral laws remains to be determined" (p. 4). Hilgard remarks that "a peculiar twist is sometimes given psychological thinking which takes the form that a process in order to be scientifically reputable must be demonstrated to occur in lower animals. If it occurs only in man, it is excused away because man possesses verbal and related abilities" (1948, p. 329). We may make the distinction between the veterinarian approach to human learning which "puts the rat into man" and what Hilgard calls the pharmaceutical approach. The latter approach tests infrahuman specimens only for their comparability with human functions. "Only if a process demonstrable in human learning can also be demonstrated in lower animals is the comparative method useful in studying it" (p. 329). Cognition theorists have more or less implicitly followed this rule. Köhler's (1925) studies on insight in animals is a good example. Tolman's position may be considered ambivalent on this score, since he "enjoys sitting on the good old phylogenetic limb" (1945). Yet, Spence (1944), Boring (1953), and others may not quite trust him up there because Tolman defines his experimental problems "phenomenologically," as attested by his well-known dictum "If I were a rat" (1938). He thus seems to "put man into the rat."

In essence, these issues add up to the broader and more fundamental question of a functional difference between animal and man. S-R theorists do not consider this difference a serious one. Thus N. Miller, in defending the equation between the learning of rats and men, uses the analogy, "the obvious differences between the bodies of Rita Hayworth and an artichoke" do not dispense with the assumption "that the same fundamental mechanism of cell-division by mitosis is involved" (1951, p. 376). By comparison, Lashley recognizes the same mechanism of forming simple associations in most organisms. Nevertheless, he maintains that what is more important are the differences in organizational functioning between lower and higher animals. "It is not the fact of learning but what is learned that differentiates animals in the evolutionary scale. The learning of higher animals involves the perception of relations which is beyond the capacity of the lower" (1949, p. 30).

CAPACITY DIFFERENCES BETWEEN MAN AND ANIMAL

It is noteworthy that biologists, medical scientists, and comparative psychologists have more emphatically emphasized the capacity distinction between man and animal than have neopositivists. Phylogenetic reductionism has been called in question by such writers as Novikoff (1945) and Goldstein (1939). Novikoff presents arguments that organic matter already constitutes a new level of organization which requires biochemical concepts and which is not reducible to molecules and atoms. He postulates that the evolutionary process proceeds through a series of new irreducible "integrative levels." Novikoff criticizes the elementarism of traditional cell theory employed in explaining the biology of multicellular organisms because it does not account for the new emergent whole of *cell organization.* Just as any organism at any point in the phyletic scale represents a new integrative level, so it is itself organized into higher and lower levels: "a knowledge of the lower levels does not enable us to predict *a priori* what will occur at a higher level. Although some may have validity for the higher level, laws of a lower level are inadequate to describe the higher level" (1945, pp. 214-215). This applies particularly to man. According to Novikoff, man's biological organism is transformed by its existence in society. Under such influence the biological becomes psychological. This is made possible by man's unique biological constitution, permitting thought, speech, and the capacity for planned action. Goldstein (1939) independently follows a similar line of thought. He challenges the view that one can benefit by exploring the functions of "lower" animals for an understanding of "higher" forms because the lower ones are presumed to be simpler. The concept of simplicity applied to lower phyla introduces more difficulties than advantages. The behavior of simpler animals turns out to be actually more obscure because it introduces *different* variables than those known to and understandable for the human investigator.

The crucial diagnostic features of homo sapiens have led G. G. Simpson to join Julian Huxley in objecting to the "nothing but" fallacy which dealing with man as an animal suggests: "It is important to realize that man is an animal, but it is even more important to realize that the essence of his unique nature lies precisely in those characteristics that are not shared with any other animal" (1950, p. 283 ff.). What are these differentiating characteristics? Herrick asserts that in man new patterns of behavior have emerged which are not known anywhere else, and he singles out the following: "Symbolic thinking is a new kind of function, though the steps by which it was fashioned can probably be traced" (1926, pp. 350-351). J. Huxley, perhaps more than any other biologist, has emphasized that every human activity, including eating, sleeping, and mating, is unique, so that "man stands alone" in nature (1927). Such biological uniqueness is, for Huxley, rooted in man's capacity for conceptual thought. This potentiality is the condition for the creation of language, the utilization of tools, and the forming and transmission of tradition. The new type of social integration which emerges on the human level is considered by Warden (1936) to be cultural rather than biosocial because of the presence of language. These arguments can be multiplied by quotations from other biologists and psychologists who have singled out the cognitive aspect of symbols as the mediating vehicle in the social behavior of man. Further, they have maintained, as Schneirla (1946) has shown, that there is a psychological difference between the pseudosocial organizations on the infrahuman level and the societal organization of man. Only human society is characterized by symbolic communication.

SYMBOLIC FUNCTION, CONCEPTS, AND PROPOSITIONAL LANGUAGE

The pertinent investigations of symbolic function bring into sharp focus the differences between the reductionist and cognitive positions. At the one pole we find philosophers such as Cassirer (1944a), who defines man as a unique "symbolic animal." He makes the capacity for representation through concepts, language, myth, and art the criteria for this uniqueness. Man no longer deals with things themselves but with self-produced symbols of things and of signs. Man alone has created symbolic systems as a new dimension of behavior.

This new reality transcends the visible signals of the physical world. It is an ideational, artificial medium which man has interposed between himself and the things of nature. At the other pole are those behavior theorists who prefer to equate symbols with signals or signs which can be learned according to the principle of substitution and stimulus generalization in conditioned response terms.

Among cognitive psychologists it was K. Bühler (1934) who most clearly formulated the problem of a "darstellende sprache" (representational speech) as different from signals and sign language. It was left, however, to the neurologists' study of speech impairment following brain injury to establish empirically the importance of representational language (Jackson, 1915; Goldstein, 1939, 1948; Head, 1926). It was Jackson who first pointed to the important difference between emotional and propositional speech — a distinction adopted by Goldstein, Head, and others. Emotional speech is confined to the expressions of subjective emotional states. Such are the vocalizations and cries of animals, the exclamatory human utterances, i.e., interjections, curse words, and expressions of affect. Propositional language exists only when speech is used to designate or describe objects and situations *independently* from a subjective state or an action context. To propositionalize means making "statements about," i.e., using language denotatively. The line of demarcation between animal and man lies in the animal's inability to propositionalize. Says Cassirer, "In all the literature on the subject there does not seem to be a single conclusive proof of the fact that any animal ever made the decisive step from subjective to objective, from affective to propositional language" (1944a, p. 30).

Language development, whether onto- or phylogenetic has been frequently "explained" as a transformation from the affective language of exclamatory interjection to the propositional language of naming and communication. Cassirer argues "that between interjection and word there is a chasm wide enough to allow us to say that the interjection is the negation of language; for interjections are employed only when one either cannot or will not speak" (1944a, p. 117). The forming of words has been explained also as a fixation of certain babbling sounds already in the repertoire of the infant via circular reflexes and rewarded imitation of heard sounds (F. L. Allport, 1924). Against this view, Goldstein cites the evidence that the baby's babbling contains all kinds of sounds, most of which never appear in later language and do not exist in the speech of adults in the child's surroundings. Therefore the child has to learn totally new sounds for words and exerts great effort in the word stage to learn some of those sounds which he previously uttered in the babbling (1948, p. 35). Goldstein's conclusions have been independently amplified and supported in a study by Lynip (1951), who used magnetic recording and a newly designed sound spectrograph. Lynip found that during the first year not a single vowel or consonant is produced that is comparable to adult vowels and consonants, and even at the end of that period the sounds of the infant are "only approximately the equivalent of an adult sound" (p. 246). This also holds for the early imitative period, even when the naïve listener has the impression of a reasonably close imitation. "The infant's vocalizations have no more relation to an adult's words than his leg-kickings have to a grown-up's genuflexions" (p. 247).

The decisive maturational step from babbling to communicative speech is the socially motivated intention to acquire and understand the word in a denotative relationship. L. A. White (1949), Lindesmith and Strauss (1949), Vigotsky (1939), and Werner (1940), among many others, have insisted that the meaning of the word is symbolic. First, the sound patterns come to represent something different from what they are as mere sounds. Second, the word is an expression of generalizing thought which culminates in genuine concept formation. Concepts are psychologically operative when the invariant relationship between the properties of an object, an action, or an idea is grasped, and when the communality of characteristics that is invariant can be abstracted from a variety of changing aspects. This process in turn makes the word a conceptual symbol. The name for an individual object in daily life does not refer to the specific uniqueness of the object; the name signifies the object as a representative of a category — an exemplification of all the possible variations allowed for by its invariant characteristics. For example, in the word "green" as applied to a given green nuance, we cognize this specific green as a representative of all possible shades of green. Workers in the field of concept formation have further emphasized that the abstraction of the invariant is a deliberate, isolating activity. A general idea is attained by *abstracting from* as well as *recognizing* the differences among the given specific variations (Werner, 1940; Hanfmann and Kasanin, 1942; Goldstein and Scheerer, 1941; Scheerer, 1949). It is obvious that we are here dealing with the most complex form of representation, which, strictly speaking, consists of a hierarchy of symbols: the sound stands for a word, the word stands for a conceptual range, and the perceptual object stands for a class (E. Heidbreder, 1945a, 1946a, b).

No discussion of word meaning should omit Helen Keller's well-known account of her sudden discovery that everything has a name. This discovery was not just hitting upon a mechanical rule, but an insight into the representational function of symbols as a general principle. Up to the crucial water-pump incident Helen had learned with difficulty to imitate the forms of the touch alphabet and to connect them with objects touched — very much as association theory would depict the learning of words as signs or cues. Following her insight, she speedily acquired many times more words than she had heretofore learned by rote, and spontaneously asked over and again for the names of objects she touched. Shortly thereafter she had the first "conscious experience of an abstract idea," namely, "think" (H. Keller, 1931, pp. 21-30).

Word meanings, however, do not stand alone. They are combined in sentence structure, which is the mature mode of communication, i.e., propositional language. According to Jackson, "The unit of speech is not the word but the proposition formed by the interaction of all the words used" (cf. Head, 1915, p. 10). The interaction is possible only when a relation among words obtains which creates a *new* meaning. And this new meaning is not based upon the mere addition of individual word meanings, but rather on their organization into a new whole. This reasoning substantiates the view that propositional language and syntactical structure are complementary. Recently Lashley (1951) made a trenchant analysis of the processes which underlie the syntactical order of language and the serial order of motor behavior. He disputes the "associative chain theory" of sequential action, according to which movements are associatively linked together so that one movement is the response to the preceding movement and the stimulus for the following movement. According to Lashley, the order of muscle innervations in speech or skilled movements (such as typing) is not a temporal sequence of direct associative connections. The pronunciation of the word "right" involves the same motor elements as the word "tire" but in reverse order. In speech such movement elements occur in so many permutations that individual associations between each of them is improbable. Correspondingly, in typing a word, finger strokes occur in all sorts of combinations, so that no single letter invariably follows another specific letter. Whether, for example, following g, "gh, ga, or gu is written, depends upon a set for a larger unit of action, the word" (1951, p. 116). The same holds for the spelling of a word, as well as for the relation of a word to the sentence. As it sounds, "the word 'right,' for example, is noun, adjective, adverb, and verb, and has four spellings and at least ten meanings" (p. 116). Its *adequate* usage and placement in a sentence will depend on the total sentence meaning and not on any temporal associations between the elements within the word or between this word and others. The temporal order must therefore be imposed upon the elements of speech by some preceding organization. It consists of the set or idea to be expressed and of a general schema which simultaneously predetermines the temporal sequence of the specific acts.

To summarize, propositional language has these characteristics. It is conceptual; it is symbolic; it is not a temporal chain of associations. The preconceived schema of the sentence as a whole determines the relational meaning between the words. Functionally, language is used to convey meaning, detached from a subjective perspective and divorced from an immediate action-situation. Language is denotative and can signify the "real" as well as the possible, and also the "unreal." S. Langer has thus captured the essence of propositional language qua "statements about":

> Animal speech never has a structure. It is merely an emotional response. Apes may greet their ration of yams with a shout of "Nga!" But they do not say "Nga" between meals. If they could *talk about* their yams instead of just saluting them they would be the most primitive man instead of the most anthropoid of beasts. They would have ideas and tell each other things, true or false . . . they would make plans and invent laws and sing their own praises as men do. (1949, p. 11)

ABSTRACT ATTITUDE

Although there are certain differences in emphasis between the theories of Cassirer, Langer, Head, and Goldstein, there is agreement among them on this point. Propositional language expresses reflective abstraction from the "immediate" by representing it through symbols.

The more these conclusions were supported by the findings on aphasics, the clearer it became that the representational processes which underlie thinking in verbal symbols extend beyond the sphere of language proper. Head (1926) himself recognized that the acts of more complex symbolic formulation require a greater degree of abstraction in their execution. His examination of aphasics disclosed that the use of language as well as thinking can undergo a characteristic change. The patients usually possess emotional speech and many speech automatisms. However, they cannot employ language intentfully in the propositional sense. A word just uttered in an action context may fail them when it is demanded outside of that situation "in the abstract." It is possible for them to ask for a cigarette or say, "It is five o'clock," but not to name subsequently a cigarette or clock upon request. When examined for his conceptual grasp of words, a patient may respond as follows: a knife presented with a fork is named "knife"; if presented with an apple, it is called an "apple parer"; with a pencil, a pencil sharpener, etc. Although he used the seemingly general word "knife," the patient actually gave an individual usage name each time, unable to say "knife" in a "categorical" sense (Goldstein, 1940). Another patient "does not understand the word 'manager' used as a generalization but can use it correctly in designating the man he worked for" (Werner, 1948, p. 274). It further turned out that cortical dysfunction, with or without aphasia, tends to impair a range of performances. The patients may be unable to demonstrate an action out of context, e.g., threatening without being angry, or to repeat untruths such as "I can write with my right hand" when the right side is paralyzed, etc. They may fail to sort objects or colors according to a guiding concept and can only group in terms of realistic use or perceptual impressions of belongingness (Goldstein, 1936, 1939).

The extensive case material has suggested two generalizations. (1) In the cortically injured, the meaning and use of language tend to be restricted to a simpler, cognitive scope of impaired conceptual thought. (2) The personality of the patient as a whole can become constricted to a primitive level. A pertinent distinction between abstract and concrete behavioral levels or attitudes has been introduced by Gelb and Goldstein (1925) and their co-workers (Goldstein and Scheerer, 1941; Hanfmann, Rickers, and Goldstein, 1944; Scheerer, Rothmann, and Goldstein, 1945) and others. According to the experimental and clinical findings of these authors, normal behavior manifests a coordination of the abstract and concrete attitudes. The performance changes in the brain-injured are symptomatic of a de-differentiation of the higher functional levels involving "abstraction." Abstract behavior detaches the self from direct "outer or inner experience," e.g., from the demand qualities of phenomenal objects and immediate thought contents. The normal person can therefore impose mediating ideational structures on the given situation, transcending its immediate realness and uniqueness. Symbolic representation, active planning, and shifting are possible. The highest degree of abstraction is the reflective awareness of the cognitive activity itself, enabling the individual to give a verbal account, e.g., of his conceptualization — which also can be formed less abstractly. E. Heidbreder (1945a,b, 1946a,b) developed a comparable theory from her extensive experiments on concept formation. She, too, makes the distinction between higher and lower forms of human cognition, e.g., the conceptual and symbolic as against the perceptual-manipulatory grasp of concrete things.

Abstraction has volitional intent whenever reflective awareness is present. We may translate this behaviorally into *self-instruction,* whereby the individual wilfully assumes a certain set corresponding to the "aufgabe." The task-execution can then take place with situational support or in a habitual fashion without step-by-step reflective guidance, until a shift is necessary. This relationship between self-instruction via abstract attitude and concrete enactment may be assumed for a wide range of actions, including, for instance, the taking on of varying roles in social behavior.

SYMBOLS, SIGNS, AND SIGNALS

Cognitive interpretations of symbolic and abstract behavior contrast sharply with two current theories, of which one reduces symbols to signs and the other tends to reduce concepts to stimulus generalization. The first is reflected in the attempts of Morris (1946), followed by Seward (1948) and Osgood (1952). The second originates directly in the Hullian school. Morris admits some difference between signal and symbol but subsumes *both* under the concept of sign. "A symbol is a sign produced by its interpreter which acts as a substitute for some other sign with which it is synonymous; all signs, not symbols, are signals" (1946, p. 25). The difference between a symbol and a signal is chiefly that the signal is provided by physical events in the environment or in the organism. The patter on the roof is a signal of rain, the heartbeat may be the signal of excitement. In contrast, the symbol is a sign provided by the organism in the absence of a signal (or object), and substituting for it. The *word* rain, for example, would substitute for the signal of rain, or rain itself. The behavioristic derivation of symbol is further established by the assertion that the sign "is a preparatory stimulus which in the absence of stimulus objects initiating response sequences of a certain behavior family causes a disposition in some organisms to respond under certain conditions by response sequences of this behavior family . . ." (p. 10). The notion of behavior family is a special case of Hull's habit family in that the response sequences terminate in similar goal objects for similar needs.

The crucial issue between Morris and cognitive theorists such as Suzanne Langer concerns the nature of the difference between signals or signs on the one hand and symbols on the other. For Langer (1951) signals and signs are synonymous and differ from symbols. Signals evoke actions but a symbol does not necessarily evoke an action appropriate to the presence of its object. The word "Napoleon" does not call out specific behavior towards Napoleon but a thought about him. "In talking about things we have conceptions of them, not the things themselves; and it is the conceptions not the things that symbols directly 'mean.' " Again, "signs announce their objects . . . whereas symbols" lead us to "conceive their objects" (Langer, p. 61). Morris, on the other hand, considers the difference between signal (or sign) and symbol one of degree. He says, "The difference is not between 'announcing' an object in the case of signals and of 'conceiving' an object in the case of symbols, but in the degree of absence or presence of the supporting conditions under which dispositions to behavior issue in overt behavior" (1946, p. 52). In reply to the example of Napoleon, Morris argues that one need not act any more in response to the word symbol of "Napoleon" than to the signalling pattern of raindrops on the roof. It suffices that the *disposition* to behavior is evoked. One wonders whether this argument would hold if the words "peace" or "democracy" or "justice" were heard instead of Napoleon. The presumed "response disposition" must then become so general and abstract that the concept becomes diluted to a behaviorist's version of an idea which can no longer be a response substitute for the behavior *per se,* be it overt or covert. Morris himself does realize certain limitations in his theory of sign response. He admits that animals seldom *produce* the signs which influence their behavior. Human language symbols "characteristically do this, and to a surprising degree. Here is a basic difference between men and animals, and until behavioral theory develops a semiotic adequate to this difference it will remain what it is today: a careful study of animals and a pious hope for a science of the human person" (p. 198).

The psychologist Osgood (1952) has extended Morris' formulations to a strictly behavioristic derivation of meaning and concept. First, Osgood adds S-R learning to Morris' notion of behavior disposition. Second, he invokes the construct of partial identity between the behavior disposition elicited by the sign and the previous behavior elicited by the object itself. "Words represent things because they produce some replica of the actual behavior towards these things as a mediation process" (p. 204). Therefore a stimulus pattern which is not the object is a sign of the object if it evokes a mediating reaction which is a fractional part of the total previously evoked behavior by the object. This mediating reaction is further defined as a self-stimulating one that leads to responses which would not occur without the *previous* association between sign and object. Productive as these propositions are, they clearly serve to rescue Morris' concept of behavior disposition from its dangerous flirtation with "idea" and confine it to the Hullian "pure stimulus act." Osgood's explanation rests upon a fractional identity of responses without really accounting for the psychological character of the identity. We have here the same situation as in some cases of problem solving, and of concept formation experiments, where the crucial feature is the *discovery* of the so-called identity. Moreover, cognition theorists may query, with G. W. Allport (1937): "How identical is an element?" To what extent does the presumed identity require the very process of abstraction which the fractional identical response is supposed to replace? There remains the further challenging psychological question of whether these self-produced signs of signs — the symbols — are ever invented without preceding object associations. Does not the language of scientific constructs, of mathematics, of art symbols, or of games such as chess testify to at least some modicum of creativity? It is difficult to decide whether the notion that the organism's responses furnish stimuli which are the "surrogates for ideas" could endow the "pure stimulus act" with original creativity. Hull (1952) has further refined the concept of "pure stimulus act," especially with regard to *human* behavior. In this respect, Hull now accepts both "molar foresight or foreknowledge" as well as purpose in Tolman's sense, and even insight (p. 150 ff.). However, the "identity of fractional antedating elements" does remain the basis of such cognitions (cf. also p. 325).

While S-R theorists such as Seward (1948) are willing to grant propositional symbols to man alone, they are at the same time eager to grant other symbols besides words to animals. Conceding that there can be no debate concerning human speech, Seward excludes this phenomenon from further discussion. Instead, he sets out to prove that animals can produce symbols and that therefore "the conative and cognitive processes of humans and other species belong on a continuum varying only in complexity" (p. 293).

Seward also interprets Morris' exposition of a symbol as a surrogate response. But it is not a behavior disposition, it is an "intermediate reaction" which substitutes for a sign by providing an internal stimulus for a response. The concession is made by Seward that even such nonverbal symbols occur in the central connections between S and R. "Stripped of all subjective implications it is a behaviorist's version of an *idea*" (1948, p. 282). Since this central process concerns something *perceptually absent* and is at the same time a *stimulus for behavior,* it would be a typical case of a distal referent in disguise. If that should be denied, the central process must be a circumlocution for an image. The generality of this image can again not be likened to a conceptual symbol. Cassirer has clearly expressed this in discussing the theory of groups:

> The concept cannot be represented in the form of an image but only in the form of a rule . . . Concepts are psychologically actualized by "schemata," not by images . . . The rule may, in simple and exact terms, be defined as that group of transformations with regard to which the variation of the particular image is considered. (1944b, p. 22).

THE COGNITIVE PROBLEM IN STIMULUS GENERALIZATION AND TRANSPOSITION

Another attempt to bring the representational character of generalizing into the orbit of S-R laws is more directly connected with the Hullian postulate of generalization. Spence (1937, 1942) made an ingenious use of this postulate in reinterpreting the configurational evidence for discrimination learning and transposition. According to Gestalt theory, transposition responses as well as many reactions to "equivalent stimuli" are governed by the *relations* between the stimuli rather than by their absolute properties. Similar to melody transposition, relational dominance is demonstrated in the type of training Yerkes and Köhler introduced. For example, a small and large box are presented and the choice of the larger rewarded. A still larger box is then paired with the rewarded and transposition tested; further stimulus pairs are added, increasing the size each step (the original size ratio is preserved and positional factors are controlled). With infrahuman subjects up to the chimpanzee the range of transposition is limited in the number of steps the test pairs can be removed from the training pair. An accepted explanation of this limit is still wanting. Lashley (1938) suggests a perceptual interpretation. Test pairs requiring far transposition may obscure the relational and emphasize the absolute character of the stimuli. He offers the analogy of a pinhead and a wagon wheel, which do not belong to the same size dimension as a dime and a quarter. Extreme stimulus sizes will have absolute properties outweighing the relational ones. For Spence (1942) this analogy lacks predictive value, since no quantitative specification can be deduced from it. He advances an alternative nonrelational hypothesis to cover both transposition and its range. Consonant with S-R tradition, he posits an individual response to each of the two absolute stimuli in the original training pair. In essence, his explanation amounts to an assumed reinforcement effect of the rewarded absolute stimulus ("excitatory potential") and an inhibitory effect of the nonrewarded absolute stimulus ("inhibitory potential"). To each of these two stimuli corresponds a generalization gradient. Since the two gradients must overlap, this will result in a differential between the excitatory and inhibitory potentials for any given pair of stimuli along a continuum of transposition (e.g., size). By mathematically constructing the curves for both gradients, Spence tries to account for the empirically found limits of transposition — confirmed with his chimpanzee subjects.

The issues in this controversy clearly carry far-reaching implications for cognitive problems, and for concept formation in particular. The presumed stimulus generalization is, of course, patterned after Hull's conditioning theory, in which Pavlov's temporary irradiation effect becomes something more permanent. In the early stages of conditioning to a tone, for example, the generalization covers a wide range, so that the gradient shows little or no slope from the point conditioned, i.e., many other tones will elicit the C-R. Only through continued reinforcement of the specific stimulus, usually by contrasting it with the "irrelevant" stimulus, will the range of generalization narrow and a gradient develop. The question is raised by Lashley and Wade whether "many instances of so-called generalization consist of nothing more than failure to observe differences" and failure to "form effective associations with specific aspects of the stimulus" (1946, pp. 82-85). Consequently, such a "generalization by default" requires the positing of relational percepts rather than absolute stimulus responses, since the latter depend on discrimination of the individual stimulus. The finer details of the Lashley-Hull controversy need not concern us here. The basic question of generalization by default seems to have been lost in the shuffle. Since it remains an unresolved problem it should invite re-examination. This need has been recognized by some S-R exponents. Mowrer and Lamoreaux (1951) have called attention to the fact that *true conditioning* may occur in one trial. The conditioning is to the *general* features of the situation as a whole. Only after further trials do responses become specific. This latter, they claim, is discrimination learning, and should not be confused with conditioning proper. Nevertheless, the tacit assumption of behavior theorists seems to be that stimulus generalization, transposition, and "genuine" generalization in concept formation are basically equivalent. This assumption may be premature if not unwarranted.

A case in point is the experimentation of M. Kuenne (1946) and Albert and Ehrenfreund (1951), who tested preschool children at progressive mental age and chronological age levels on a size transposition problem. The experimenters adopted Spence's gradient of generalization for predicting the range of transposition. The more advanced groups succeeded significantly better on a far transposition than the less mature groups. The investigators noted that the successful children were those who spontaneously used "verbal cues" ("it's always the smaller one"). It is significant that the transposition results of the nonverbal and younger children were consonant with Spence's gradients. The gradient for the verbal children, on the other hand, remained in principle horizontal, indicating complete transposition throughout. This deviation from the Spence gradients is attributed to verbal cues, which facilitate the far transposition. The psychological nature of these verbal cues is not further elucidated, but the term "verbal cues" bespeaks an associationistic conception of language acquisition.

A hypothesis of lower and higher processes to cover stimulus generalization, limited transposition range, and the findings on far transposition has been offered by Scheerer (1949). For the lower processes he posits, in the same vein as Lashley, that in generalization by default there is a lack of discrimination, a generality of response based on inarticulate differentiation. For example, in early conditioning the response is made to sound, brightness, form, i.e., to an unanalyzed whole impression against a homogeneous ground. "Since no other specific attributes are herein discriminated, the same response can occur within this undifferentiated range of stimuli" (1949, p. 249). Perceptually, other sounds, brightnesses, or forms have the character of likeness or sameness at this stage. Correspondingly, in limited transposition, the differential characteristics between the original pair and the new pair are not cognitively registered, but the preserved relations of brightness or size lead to a concrete impression of likeness. Köhler, in a personal communication to Wapner, indicated that transposition alone does not constitute evidence for insight. Köhler doubts whether "an animal who chooses relationally and therefore tends to transpose, need be aware of the fact that his choice is relative" (Wapner, 1944, p. 51).

Such an impression of likeness may be illustrated on the human level by the fact that most people do not hear any difference when a melody is transposed, up to a certain point; " 'it's just the same' to them — there are no differentials for them from which to generalize" (Scheerer, 1949, p. 249). This also holds for many other Gestalt transpositions, e.g., triangular or square forms, where the total impression of pointedness or angularity may persist without cognitive discrimination of differences. A generality such as this, based on global likeness, underlies the so-called stimulus generalization as well as the limited range of tested transposition. This likeness will break down when the perceptual disparity increases so that other absolute stimulus characteristics become dominant. It is at this point that higher-order cognitive processes must take over if the factor of similarity or relational likeness is to be abstracted. This process is conceptual in kind and qualitatively different from concrete responses to perceptual impressions of likeness. In the light of this interpretation the difference between the verbal and nonverbal groups in the Kuenne-type experiment may become explicable. The verbal cues are word meanings, expressing a principle.

Spence is right in demanding quantitative predictions as to *when* the presumed global likeness breaks down. The fact that this empirical question has not yet been answered satisfactorily does not decide the superiority of absolute over relational hypotheses. A relational theory can well account for primary stimulus generalization, i.e., the period before the conditioned stimulus becomes individuated by the reinforcement; it does account for the occurrence of primitive discrimination in terms of patterned percepts, e.g., figure-ground; and it accounts for the shift from near transposition to far transposition by distinguishing between pseudo and true generalization. In another connection, Maier suggests that the transfer of responses to new situations may depend upon either one of the following processes: "(a) reactions to likeness alone, and (b) abstraction, which involves a balance between reactions to likenesses and differences" (1941, p. 179). This statement is consonant with the point of view presented here, that in true generalization the differences as well as the communalities are discriminated, and the apparent perceptual disparities are subordinated to a common principle. Numerous experimental findings on concept formation and on the acquisition of word meanings support this position (Hanfmann and Kasanin, 1942; Scheerer, 1949; Werner and Kaplan, 1950). The conclusions of these authors all point in a similar direction, namely, to a distinction between reactions governed by total impressions which are not analyzed and processes of analytic generalizing abstraction.

Recently Gibson (1951) has independently taken an almost identical position. In his discussion of schematic perception he distinguishes a preconceptual or preverbal "stimulus generalization" from genuine conceptual thinking in verbal symbols. He succinctly sums up the decisive developmental differences: "As primary stimulus generalization decreases, a different kind of generalization increases . . . human reactions manifest not only a kind of generalization based on ignorance but also a kind of generalization based on knowledge" (1951, pp. 107-108). Scheerer (1953) has extended the notion of indiscriminate, primitive generalization to incidental learning in personality functioning. Acquired coping mechanisms of the individual can often be traced to an undifferentiated cognition which structures aspects of certain situations as the same without an awareness of differences (e.g., fears, personal likes and dislikes, general means-ends patterns).

Maier (1941) and Wapner (1944) have likewise questioned the premise that true generalization occurs in the more primitive type of equivalence and transpositional reactions. Wapner's experiments proved that partially decorticated rats show *more* relational responses than do intact animals. "This precludes any possibility of interpreting the relational aspects of equivalent behavior as a higher mental process" (1944, p. 49). He concludes that the term "generalization" makes for a confusion of higher mental processes with primitive types of perceptual reactions. Along with Werner (1940) and Goldstein and Scheerer (1941), Wapner argues that the end result of a performance may in one case stem from the process of abstraction while in another the same result may occur on a concrete perceptual level. Achievement and process therefore must be separately analyzed (cf. Scheerer, 1946). The principle of differentiating achievement and process has been implicitly or explicitly adopted by most cognition theorists, with the possible exception of Brunswik. Their emphasis upon the mediating processes in cognitive representation allows for qualitative distinctions as to the nature of both mediation and cognition. In contrast, S-R theorists, even when they postulate intermediary processes, have attempted to conceptualize them as uniform. Therefore qualitative differences in function according to levels do not play any significant role. The S-R constructs of mediation, as far as they are offered, are placed on a continuum from simple to complex, differing only quantitatively.

COGNITIVE DEVELOPMENT

It is an apparent paradox that those who are most inclined towards genetic reductionism are at the same time those who are least sensitive to or cognizant of the problems of developmental change. It is for this reason that cognition is neither given a special place in development nor is it treated developmentally, in terms of genetic stages. Investigators concerned with such questions have been critical of an additive conception of development which the terms "simple" and "complex" often imply. Goldstein (1936) marshalled data which show that the criterion of difficulty of performance cuts across a numerical scale of simple and complex. The lifting of a single finger is more difficult than the extension of the whole hand. It requires the voluntary isolation of the digit. It is easier to "scratch" one's nose or hit an object with a finger than it is to point to the nose or the object with the same finger in the act of "showing where." The latter requires a symbolic function. Typically, the pathologically concretized person may reproduce from memory a house composed of many sticks. Yet he fails to reproduce correctly the position of a single stick which is not tied to a concrete object. This requires the grasp of spatial directions in the abstract. The implied hierarchy of functional organization calls for a new criterion of simple and complex that is not merely additive. Such a criterion is the *degree of differentiation* which is the precondition for qualitatively different levels of performance. The explanatory fruitfulness of this construct for explaining functional differences has been attested in biology as well as in developmental psychology (Weiss, 1939; Child, 1924; W. Stern, 1930; C. H. Bühler, 1930; Werner, 1940; and Piaget, 1937).

Perhaps the most systematically formulated theories of mental development have been those of Piaget and Werner. We are singling these out because both made the emergence of cognitive processes the focus of their research. Both have gathered their material independently, and neither is reductionistic. Werner's earlier formulations on this point are readily matched by the recent statement of Piaget: "It is not a matter of reducing the higher to the lower . . ." The task is rather to "establish a reciprocal relationship between the two terms which does not at all destroy the originality of the higher term" (1952, p. 8; cf. Werner, 1948, pp. 5, 22). The genetic interrelationship between lower and higher levels is conceived by both as a progressive sequence from the lower to the higher. Each new level has an organizational structure of its own which forms a stage. Thus each level is a "whole" in terms of functional organization. It emerges from the preceding one but differs from it; what was a whole on the lower is now integrated as a subordinate part of the higher level. *The genetic conception here advanced supports and complements the hierarchy of mental functions previously discussed.* In this connection the groundwork laid by Piaget and Werner must be considered in any theory of cognition because they endeavor to trace the characteristic representational form and content of each stage as well as the course of cognitive development as a whole. Despite differences in terminology many of their conclusions are mutually translatable without major distortion. In view of this congruency in their thinking we shall select portions of their research so that they will supplement each other.

Piaget's importance for cognitive theory lies in the following keynotes of his system. There is first his explicit postulate of cognitive structures as mediating processes in perception, acting, and thinking. Distal achievement, in contrast to Brunswik, is studied in terms of the particular intervening organismic activities. Second, Piaget recognizes these mediating processes chiefly as representations. Third, he implements this conception by genetically tracing the emergence of representations through ontogenetic stages. He follows the intellectual evolution from the cradle to maturity by empirical samplings at each stage. This development is treated in a framework of biological principles.

Mental processes grow out of basic organismic processes which lead to cognitive structuring in accordance with the *adaptation* of the organism at a given stage. Adaptation consists of an interplay between assimilation and accommodation tending towards equilibrium. Physically, assimilation means that the organism absorbs substances and changes them into something compatible with its own substance. Behaviorally it means that the organism, instead of submitting passively to the environment, modifies it by *imposing* certain structures of its own on it. Accommodation signifies the action of the environment on the organism, which becomes affected or changed accordingly. However, the outer stimulations as such cannot have a direct impact. They can only *modify* the process of assimilation. Translated behaviorally, this model can be exemplified in sensory-motor coordination. Manipulation contains the factor of assimilation in grasping the object, and the factor of accommodation in adjusting the prehension to the form and resistance of the object grasped. With regard to personality functioning, the same principle holds for the accommodation of the individual to the social influences and pressures of his environment, e.g., norms, conventional ideas, and "obligatory truths" (1950, p. 159). The model also implies a *changed end-stage* in the person. This is typified when in vision a peripheral object is fixated. The eye movements change the initial state of lacking clarity to an end-stage of optimal clarity — the object is "assimilated," while accommodation to it takes place simultaneously. Correspondingly, the assimilation of information or new ideas and goal attainments results in a change of the *person*. This important aspect of personality development has been emphasized by G. Murphy (1947) and by G. W. Allport (1937), by Goldstein in the notion of enrichment, and by Schilder (1942) in the theory of completion states.

The equilibria involved in these adaptive cycles are considered quasi-stable systems. They have a notable similarity to current theories of self-regulation, feedback, and to the reflex scheme proposed by Koffka (1935). For Piaget, the adaptive equilibria in mental functions differ for each developmental stage, so that general adaptation is never complete but continues. Such stages as "sensory-motor intelligence," "preconceptual," and "intuitive thought" are quasi-equilibria, each at a specific level of cognitive development. The relative incompleteness of adaptation at each stage may be equated to immaturity, i.e., to the level of maturation at which the balance between assimilation and accommodation is excessively tipped in either direction. Methodologically speaking, Piaget designs his experiments and observational procedures so as to detect the errors in the child's performance on tasks and to analyze the processes which underlie the solution or nonsolution. In principle, the errors are conceptualized as the inability to "decenter," and the developmental progress is conceptualized as an *increased ability to decenter.* The meaning of "decentering" may be appreciated by first thinking of the tendency in the young child to become "centered" or rigid. The nine-month-old baby can reach for a milk bottle hidden before his eyes under one of two pillows. If then the bottle is hidden, in his view, under the other pillow, the baby will repeat reaching for it under the first pillow. In other words, certain aspects of a percept, an action sequence, a situational perspective from one point of view can become so dominant that a child is "fixated" and cannot shift orientation. Only by such a shift, i.e., by decentering, can the present cognitive structure be changed and widened to include other and new perspectives. (The similarities to Goldstein's concept of ability to shift and Wertheimer's recentering are obvious.)

The progressive decentering which Piaget describes in his analysis of stages may be systematized as follows:

I. *Sensori-motor intelligence* (birth to 1½ years). Reflexes, e.g., grasping and sucking, are the earliest "perceptual-motor schemata" (1950, p. 100; 1952). Through decentering they come to form "parts" of primitive action wholes or "habits." These simple schemata are progressively decentered and integrated into new complex units, those of voluntary action and means-end behavior. Thus grasping is used first to produce rattling when a string is pulled, later to remove a screen which partially hides an anticipated goal, finally to obtain absent objects. The last act marks the beginning of representation and the notion of external permanence. Delayed action, detouring, and imitation appear. The functional character of things governs this stage. Facts are not noted for their own sake but only in relation to needs (autism). Make-believe play is restricted to using one's "body to produce an action foreign to the present context," as in feigning sleep.

II. *Preconceptual thought* (1½ to 4 years). The schemata of make-believe play become real symbols when dissociated from action with one's *own* body, e.g., putting a doll to bed. This "representative imitation" coincides with speech, marking the growth of a general "symbolic function" which encompasses "symbolic play, imaginal representation, and verbal thought" (1950, p. 126). Language is still tied to action and desires, as in naming. Names signify discovered action aspects of the objects and become part of them. Since the child does not invent language but is exposed to it, he has to "accommodate" to the existing system of conventional meanings without fully understanding them. Thus he assimilates word meanings in accordance with the egocentrism of this stage. Piaget's illustrations of preconceptual thought parallel those of W. Stern (1930). "'Bow-Wow' is said to a dog, to a hen, a cowbell, cows, guinea pigs, and a cat; subsequently to anything moving from an ant to a tractor in the field" (1951, p. 218). In contrast to concepts, "in a schema such as 'bow-wow,' there is merely a subjective feeling of kinship between the related objects" (1951, p. 220). Such "participations" figure prominently in causal attributions as well. A mountain changes appearance as the child travels around it. Unable to integrate the varying perspectives, he claims the mountain is changing. Though the idea of permanent objects has been formed in the immediate action sphere, this is not so "where distant space and reappearances at intervals are concerned" (1950, p. 128). The child "will be just as likely to say 'slug' as 'slugs' . . . without deciding whether the slugs encountered in the course of a single walk are one individual, a single slug, or a class of distinct individuals" (1950, p. 127; 1951, p. 225).

III. *Intuitive thought* (4 to 7 years). The use of representations for reasoning is limited by perceptual dominance. When faced with the problem of logically integrating several aspects of the same object or the same aspect of different objects, the child cannot disregard that aspect which is perceptually obtrusive. Some typical experiments illustrate this (1950, pp. 129-135). (1) The child puts an equal number of beads into each of two vials identical in shape and size. He also acknowledges the equal amount. Next, vial *A* is emptied into a taller and thinner glass, while vial *B* remains unchanged in sight. The child now states that the number of beads has changed, even though he avers that none have been added or taken away. Some children say that the new glass contains *more* beads because it is "higher," others insist that there are *fewer* beads because it is "thinner." Their reasoning is centered on what is to them the prominent perceptual aspect, height or width. Unable to decenter either aspect, they cannot grasp the mutual interdependence of height and width and abstract the conservation of quantity. Similarly, they fail to group two different classes of objects in one superordinate class when the two subclasses are *perceptually* different. (2) Three colored beads are strung on a rigid wire partially enclosed by a tube. As the beads are pushed through the tube, the child is asked to predict the order in which they will emerge, and also their order upon coming back through the tube. At the preconceptual stage, the original order (*ABC*) could be predicted but not the reverse order (*CBA*), which is *now* also mastered. However, when the experimenter rotates the wire stepwise by 180 degrees (½ turn), the child encounters difficulties. He can discover that a half-turn leads to a reversal and a full turn "repeats" the original order, but he cannot predict the order for many rotations. The reason for success or failure is that the child can represent to himself through action in thought the order of the beads for a *few* turns of the wire but cannot decenter his perceptually dependent imaginary manipulations in order to derive the general principle applicable for any number of turns.

IV. *Concrete operations* (8 to 11 years). In contrast to common usage, concrete operations are actions that have become abstract. They are manipulations in thought, based on inductive abstraction from below which overcomes the "intuitive" dependency on action context and situations. In the rotating-wire problem the child can now discover the rule that all half-turns reverse the order; all full turns repeat the original order. He thus predicts any number of turns. Concrete operations are, however, limited to manipulating "ideas of objects" (1950, p. 139). Logical form is still dependent on its concrete content. In the vial problem, the child can now abstract the conservation of *quantity*, being able to decenter height or width and combine them into a logical structure. But he cannot yet conceive the conservation of *volume* of two identical pieces of dough when one is changed in shape. At this stage the perceptual potencies of height, width, and length still hinder a balanced abstraction so that these aspects cannot be conceptually integrated. Correspondingly, the child cannot deduce the identical principle or concept which he discovered and utilized with *situational* support when the same problem is presented merely in verbal form (1950, pp. 146-149).

V. *Formal thought* (12 years on). The child can manipulate ideas about ideas or, as Piaget puts it, "perform operations upon operations." It is the level of abstraction at which reflective reasoning and formal logical thinking become possible. It permits mobility and reversibility of operations; the ability to conceive of the possible and the impossible; the formulation of propositions about space, time, number, as systems; and the carrying out of hypothetico-deductive reasoning which proceeds from assumptions not necessarily related to reality but related by logical necessity, independent of evidence from experience. Piaget emphasizes the attainment of the notion of "conservation" through "grouping operations." Conservation is the attribution of permanence to objects and to relations in perception and thinking. It develops progressively from the perceptual constancies to abstract concepts. For example, conservation of quantity, weight, volume, etc., are shown to differ at various stages. Such attributions of permanence demand the decentering processes which have here been traced.

We have epitomized the characteristics of these stages because Piaget's latest formulations in the English language provide the systematic basis for the understanding of his former theories on causality, language, moral judgment, etc. It becomes clear that the given age limits are not absolute but only reference points for the sequence of development. The stages are furthermore combined into larger, more inclusive phases which have a general cognitive and ego organization in common. They are the phases of autism and egocentrism, subject of much speculation. For Piaget, autism is the *assimilation of reality* to organic needs, mainly in the sensori-motor stage. A decentering occurs when the need-satisfying "autistic symbolism" becomes transformed into representations of reality, i.e., when the child increasingly accommodates to the environment.

Egocentrism, the next phase, is defined as a lack of reversibility in operations involving cognitive and social perspectives. Piaget himself concedes that the term "egocentrism" is an unfortunate choice (1951, p. 285). It does not signify self-centeredness but "a quite unconscious and natural illusion of perspective which precedes moral egoism and conscious egocentricity" (1937, p. 48). In direct experience, the child is not conscious of a distinction between subjective and objective, hence of his own "subjectivity." He has no awareness of the private or nonprivate character of a point of view and therefore cannot differentiate between what might be his own or another's frame of reference (1950, p. 160 ff.; 1951, p. 290). Since a frame of reference is necessary for orientation, a child may have the perspective of another without experiencing it as such. Young children often imitate without realizing it, believing they originated the action, just as they may confuse their own experience with that of another. The adaptive balance between accommodation and assimilation in egocentrism makes for a stepwise differentiation between self, objects, and persons. It culminates when persons are seen to have perspectives of their *own* (i.e., cognitive representations).

In evaluating Piaget's work it must be recognized that he is stronger on theory and ingenious experimentation than on rigorously controlled methods. The criticisms of some of his earlier findings, especially by Susan Isaacs (1930), Dorothea McCarthy (1930), and others, are well known. The chief objections have been: (1) Piaget treats primarily of the rational elements in child development and neglects the affective and conative dimensions. (2) Piaget does not give sufficient weight to learning, experience, and socio-environmental factors. (3) His "stages" are neither distinct nor unitary functional levels. They overlap chronologically. They distribute over different performance areas at the same age. Actually the process of development is diversified as well as gradual, but not in evolutionary steps. It is conceivable that the critical reserve in accepting Piaget's theses will be lessened after his more recent work has been absorbed into current psychology. The documentation in this later phase appears to be more exact, although again a larger sampling is missing. Dennis, after citing chapter and verse of Piaget's "sins" in methodology, adds to his critique this comment:

> . . . Piaget can neither prove nor disprove his hypotheses. That must be done by others . . . We turn to Piaget for ideas, not for statistics . . . Following an unfavorable first reaction to his books, the writer attempted to show how outrageous were some of his views, and, by gathering data with the expectation of refuting his theories, became convinced that many of them were sound." (1951, p. 161)

The assumption of developmental levels is also shared by Werner. In clarifying misconceptions of the term, "developmental level," Werner and Kaplan state: "the stage of maturity characterizing an individual or a group must not be conceived statically, as a fixed pattern of operation; it should rather be understood dynamically, as a potential range of genetically graded functions. Mental growth is definable in terms of reorganization due to the inclusion of new forms of operations. This implies that an individual, depending on outer or inner circumstances, may operate at genetically differing levels" — i.e., depending on the difficulty of the tasks (1950, p. 101). Werner includes the emotional dimension in his genetic scheme, in accord with his holistic point of departure. He derives general laws of development from a wide range of empirical data in comparative biology, psychology, anthropology, and psychopathology. At the base of Werner's treatise is the formal principle that every genetic sequence has an intrinsically required order and direction. This developmental psychology does not deal with actual historic succession but rather with the functional conditions for primitive and advanced forms of cognition. A statement from Goethe's morphology, cited by Werner, is apropos:

> The more perfect the creature becomes, the less similar become the parts to one another. On the one genetic pole the whole is more or less similar to the parts, and on the other the whole is dissimilar to the parts. The more nearly equal the parts, the less are they subordinated one to another. Subordination of the parts indicates the perfect creature. (Werner, 1948, p. 40)

Developmental direction is thus conceptualized as a movement from an undifferentiated stage towards an increasing differentiation of parts; it goes hand in hand with a progressive hierarchy of function. In this hierarchy the foci of dominance shift at different genetic levels. At one level the dominant center is a perceptual-motor affective orientation, e.g., towards expressive features of things and persons. At a later stage these reactions become integrated with and subordinated to a new center of dominance, i.e., cognitive-conceptual control. Development is analyzed into four aspects, which reflect lower and higher degrees of differentiation:

(1) FORMAL ASPECTS: *Diffuse-articulate.* Lack of articulation into distinct parts makes for *diffuseness*, e.g., in primitive perception and thought. Another example is noise, or a glissando in music. In contrast a melody has *articulation* because the parts constituting the whole are determinable.

(2) DYNAMIC ASPECTS: *Rigid-flexible; labile-stable.* Rigidity is a lack of variability in response when a shift is called for. In children this lack of flexibility may be typified by their sensitivity to change, their tendency to all-or-none reactions, their dependency on definite order, e.g., on eating and dressing sequences and ritualistic modes of action. *Lability* is expressed in the fluidity of behavior, e.g., the distractibility of young children, their impulsivity and the transient character of their emotions. *Stability* as well as *flexibility* stem from an increased central control of the direction of behavior. Both "preserve the functional equilibrium in the face of mutable situations" in contrast to lability and rigidity, in which the responses are more passively governed by the momentary valences (1948, p. 55).

(3) COGNITIVE-AFFECTIVE ASPECTS: *Syncretic-discrete.* Syncretic signifies an indistinctness between the experiences of inner states and of objects; a lack of separation between self, other, and environment; a lack of segregation in personality regions; a lack of demarcation of reality levels. Cognitively, all this is a merging of otherwise separable experiences into a global meaning. So-called "pars pro toto" thinking in magic is an instance of syncretism. Syncretic factors enter into such child notions as the following: the child asks his mother for fish, because eating it will make him a good swimmer; having had a dream, he believes that it happened in reality; having "wished" a wish, he believes he has communicated it. On the affective side, syncretism expresses itself in reacting to the "physiognomic" expressive features. Persons and things are threatening, inviting, sad, gay, angry, friendly. On the action side, syncretism manifests itself in responding to objects as "things of action." They are chewable, graspable, throwable, hideable, bounceable. At this stage the various sense modalities and the affective motor sphere are intimately interwoven. Such a "total participation" fosters syncretic behavior and cognition. This recedes with greater individuation and *discreteness* of mental contents and functions.

(4) IDEATIONAL ASPECTS: *Concrete and abstract.* These thought forms are implied in all phases of differentiation, and Werner refers to them in describing levels of conception such as space, time, causality, etc. In particular, he traces the genesis of abstraction from its primitive preverbal forms to the symbolic conceptual. For Werner, true generalization is preceded by "analogous functions" so that at each genetic level similar achievements can be attained via different processes (1937). Relations are first realized through sensori-motor activities, e.g., in the transpositional use of action-things; then through perceptual

groupings when sameness and difference are concretely apprehended; later through quasi-class concepts in terms of likeness. The next step is representation through schemata which culminate in genuine concepts. The same principle is applied to the stages of naming: the word expresses the physiognomic aspects of the object; the word represents the object schematically; and the word is a conceptual symbol of a class. The achievement of naming objects is mediated through different processes on each of these steps, even though there is the "same" effect of naming. This is also demonstrated in the realm of number development. What is accomplished concretely on the lower level is attained abstractly on the higher one. It is Werner's general thesis that such analogous functions of a lower order must emerge before a function of the next higher order can develop.

The above resumé of Piaget and Werner should speak for itself with regard to the cogency of their contributions for cognitive theory. An extensive analysis of the correspondences and differences in their systems would be beyond the scope of this chapter. Two main parallels should be stressed, however. The principle of differentiation and hierarchization matches the principle of decentering. The concept of syncretism corresponds largely to that of egocentrism, and the subaspects of both cognitive modes are complementary in many respects. Finally, the entire developmental dimension from concrete to abstract is treated in essentially similar form. Werner speaks of analogous functions and Piaget of grouping schemata as the functional equivalent of concepts.

SUMMARY

We have outlined some of the current thinking on cognition. The major trends were seen to converge on the problem of representation. Cognition was identified as a centrally mediated process of representing external and internal events. The functions of phenomenal and of silent organization in perceiving, thinking, and learning were examined in their significance for cognition. An analysis of cognitive processes was presented in terms of different levels of organization. The implications of the "levels concept" were sketched with regard to generalization, concept formation, and propositional language. This was supplemented by a consideration of pertinent developmental theories of cognition. The role of cognitive factors in motivation, attitude, and emotion was explored and related to relevant interpretations of ego-function and of the self concept.

A special attempt was made to delineate some basic tenets of cognitive theory within the context of contemporary psychological issues. In this framework we contrasted two main positions: Organizational versus Stimulus-Response orientation. We have emphasized the communalities rather than the differences among the cognitive as well as among the S-R theorists. This was done in order to clarify the critical issues through a problem-centered approach as against a descriptive comparison of individual systems. Therefore only selected problem areas were discussed to illustrate these issues. These were the areas of molar-molecular, conation and cognition, cognitive development and genetic reductionism.

References

Adams, D. K. The organs of perception: Sentiments. *J. Pers.*, 1953, *22*, 52-59.

Albert, E., & Ehrenfreund, D. Transposition in children as a function of age. *J. exper. Psychol.*, 1951, *41*, 30-38.

Allport, F. H. *Social psychology*. Boston: Houghton Mifflin, 1924.

Allport, G. W. Attitudes. In C. Murchison (Ed.), *A handbook of social psychology*. Worcester: Clark Univ. Press, 1935. Pp. 798-844.

Allport, G. W. *Personality: A psychological interpretation*. New York: Holt, 1937.

Allport, G. W. The ego in contemporary psychology. *Psychol. Rev.*, 1943, *50*, 451-478.

Allport, G. W. The psychology of participation. *Psychol. Rev.*, 1945, *52*, 117-132.

Allport, G. W. Effect: A secondary principle of learning. *Psychol. Rev.*, 1946a, *53*, 335-347.

Allport, G. W. Personalistic psychology as science: A reply. *Psychol. Rev.*, 1946b, *53*, 132-135.

Allport, G. W. Scientific models and human morals. *Psychol. Rev.*, 1947, *54*, 182-192.

Alpert, A. The solving of problem situations by preschool children. *Teach. Coll. Contrib. Educ.*, 1928, No. 323.

Ames, A., Jr. Visual perception and the rotating trapezoid window. *Psychol. Monogr.*, 1951, *65*, Whole No. 324.

Angyall, A. *Foundations for a science of personality*. New York: The Commonwealth Fund, 1941.

Asch, S. E. Forming impressions of personality. *J. abnorm. soc. Psychol.*, 1946, *41*, 258-209.

Asch, S. E. *Social psychology*. New York: Prentice-Hall, 1952.

Bartlett, F. C. *Remembering*. London: Cambridge Univer. Press, 1932.

Bateson, G. Information and codification: A philosophical approach. In J. Ruesch & G. Bateson, *Communication: The social matrix of psychiatry*. New York: Norton, 1951. Pp. 168-211.

Birch, H. G. The relation of previous experience to insightful problem-solving. *J. comp. Psychol.*, 1945, *38*, 367-383.

Birch, H. G., & Bitterman, M. E. Sensory integration and cognitive theory. *Psychol. Rev.*, 1951, *58*, 355-361.

Boring, E. G. The use of operational definitions in science. *Psychol. Rev.*, 1945, *52*, 243-245.

Boring, E. G. Mind and mechanism. *Amer. J. Psychol.*, 1946, *59*, 173-192.

Boring, E. G. Visual perception as invariance. *Psychol. Rev.*, 1952, *59*, 141-148.

Boring, E. G. A history of introspection. *Psychol. Bull.*, 1953, *50*, 169-189.

Bridgman, P. W. Science: Public or private? *Phil. Sci.*, 1940, *7*, 36-48.

Bridgman, P. W. Rejoinders and second thoughts. *Psychol. Rev.*, 1945, *52*, 281-284.

Brown, J. F. The visual perception of velocity. *Psychol. Forsch.*, 1931, *14*, 199-232.

Bruner, J. S. Personality dynamics and the process of perceiving. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: An approach to personality*. New York: Ronald Press, 1951. Pp. 121-147.

Bruner, J. S., & Postman, L. On the perception of incongruity: A paradigm. *J. Pers.*, 1949, *18*, 206-

Bridgman, P. W. Science: Public or private? *Phil.*

223.

Brunswik, E. Psychology in terms of objects. In H. W. Hill (Ed.), *Proc. 25th Ann. Celeb. Inaug. Grad. Stud.*, Univer. S. Calif., 1936. Pp. 122-126.

Brunswik, E. Points of view: Components of psychological theorizing. In P. L. Harriman (Ed.), *Encyclopedia of Psychology*. New York: Philosophical Library, 1946. Pp. 523-537.

Brunswik, E. Discussion: Remarks on functionalism in perception. *J. Pers.*, 1949, *18*, 56-65.

Brunswik, E. The conceptual framework of psychology. *International Encyclopedia of Unified Science*, Vol. I, No. 10. Chicago: Univ. of Chicago Press, 1952.

Bühler, Charlotte. *The first year of life*. New York: Day, 1930.

Bühler, K. *Sprachteorie: Die Darstellungsfunktion der Sprache*. Jena: Fischer, 1934.

Buytendijk, F. J. J. The phenomenological approach to the problem of feelings and emotions. In M. L. Reymert (Ed.), *Feelings and emotions*. New York: McGraw-Hill, 1950. Pp. 127-141.

Cannon, W. B. *The wisdom of the body*. New York: Norton, 1939.

Cassirer, E. *An essay on man*. New Haven: Yale Univ. Press, 1944a.

Cassirer, E. The concept of group and the theory of perception. *Philos. and Phenom. Res.*, 1944b, *5*, 1-35.

Chein, I. Behavior theory and the behavior of attitudes: Some critical comments. *Psychol. Rev.*, 1948, *55*, 175-188.

Child, C. M. *Physiological foundations of behavior*. New York: Holt, 1924.

Claparède, E. Recognition et moiité. *Arch. de Psychol.*, 1911, *11*, 79-90. In D. Rapaport (Ed.), *Organization and pathology of thought*. New York: Columbia University Press, 1951. Pp. 58-75.

Dennis, W. D. Cultural and developmental factors in perception. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: An approach to personality*. New York: Ronald Press, 1951. Pp. 148-169.

Doob, L. W. The behavior of attitudes. *Psychol. Rev.*, 1947, *54*, 135-156.

Duncker, K. Uber induzierte Bewegung. Ein Beitrag zur Theorie optisch wahrgenommener Bewegung. *Psychol. Forsch.*, 1929, *12*, 180-259.

Duncker, K. On problem solving. *Psychol. Monogr.*, 1945, *58*, 1-112.

Feigl, H. Operationism and scientific method. *Psychol. Rev.*, 1945, *52*, 250-259.
Feigl, H. Principles and problems of theory construction in psychology. In W. Dennis (Ed.), *Current trends in psychological theory.* Pittsburgh: University of Pittsburgh Press, 1951. Pp. 179-213.
Fromm, E. *Man for himself.* New York: Rinehart, 1947.

Gelb, A., & Goldstein, K. Psychologische Analysen hirnpathologisches Fälle. X. Über Farbennamenamnesie. *Psychol. Forsch.*, 1925, *6*, 127-186.
Gibson, J. J. *The perception of the visual world.* Boston: Houghton Mifflin, 1950.
Gibson, J. J. Theories of perception. In W. Dennis (Ed.), *Current trends in psychological theory.* Pittsburgh: University of Pittsburgh Press, 1951. Pp. 85-110.
Goldstein, K. The problem of the meaning of words based upon observation of aphasic patients. *J. Psychol.*, 1936, *2*, 301-316.
Goldstein, K. *The organism.* New York: American Book Company, 1939.
Goldstein, K. *Human nature.* Cambridge: Harvard University Press, 1940.
Goldstein, K. *Language and language disturbances.* New York: Grune and Stratton, 1948.
Goldstein, K. On emotions: Considerations from the organismic point of view. *J. Psychol.*, 1951, *31*, 27-50.
Goldstein, K., & Scheerer, M. Abstract and concrete behavior: An experimental study with special tests. *Psychol. Monogr.*, 1941, *53*, No. 2.

Hanfmann, Eugenia, & Kasanin, J. Conceptual thinking in schizophrenia. *Nerv. Ment. Dis. Monog.*, 1942, *67*.
Hanfmann, Eugenia, Rickers-Ovsiankina, Maria, & Goldstein, K. Case Lanuti: Extreme concretization of behavior due to damage of the brain cortex. *Psychol. Monogr.*, 1944, *57*, No. 4.
Harlow, H. F. Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *J. comp. physiol. Psychol.*, 1950, *43*, 289-294.
Harlow, H. F. Mice, monkeys, men, and motives. *Psychol. Rev.*, 1953, *60*, 23-32.
Hartmann, H. Ego psychology and the problem of adaptation. In D. Rapaport (Ed.), *Organization and pathology of thought.* New York: Columbia University Press, 1951. Pp. 362-396.
Head, H. Hughlings Jackson on aphasia and kindred affections of speech. *Brain*, 1915, *38*, 1-27.
Head, H. *Studies in neurology.* London: Oxford Univer. Press, 1920.
Head, H. *Aphasia and kindred disorders of speech.* New York: Macmillan, 1926.
Hebb, D. O. *Organization of behavior: A neuropsychological theory.* New York: Wiley, 1949.
Heidbreder, Edna Toward a dynamic psychology of cognition. *Psychol. Rev.*, 1945a, *52*, 1-22.
Heidbreder, Edna The attainment of concepts — A psychological interpretation. *Trans. N. Y. Acad. Sci.*, 1945b, *7*, 171-188.
Heidbreder, Edna The attainment of concepts. I. *J. gen. Psychol.*, 1946a, *35*, 173-189.
Heidbreder, Edna The attainment of concepts. II. *J. gen. Psychol.*, 1946b, *35*, 191-223.
Heider, F. Ding und Medium. *Symposium*, 1927, *1*, 109-157.
Heider, F. Environmental determinants in psychological theories. *Psychol. Rev.*, 1939, *46*, 383-410.
Heider, F. Social perception and phenomenal causality. *Psychol. Rev.*, 1944, *51*, 358-374.
Heider, F. Unpublished manuscript, 1953.
Heider, F., & Simmel, Marianne A study of apparent behavior. *Amer. J. Psychol.*, 1944, *57*, 243-259.
Helson, H. Adaptation-level as frame of reference for prediction of psycho-physical data. *Amer. J. Psychol.*, 1947, *60*, 1-29.
Helson, H. Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 1948, *55*, 297-313.
Herrick, C. J. *Brains of rats and men.* Chicago: Chicago Univer. Press, 1926.
Hilgard, E. R. *Theories of learning.* New York: Appleton-Century-Crofts, 1948.
Hilgard, E. R. The role of learning in perception. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: An approach to personality.* New York: Ronald Press, 1951. Pp. 95-120.
Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
Hull, C. L. The discrimination of stimulus configurations and the hypothesis of afferent neural interaction. *Psychol. Rev.*, 1945, *52*, 133-142.
Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
Huxley, J. *Man stands alone.* New York: Harper, 1927.

Isaacs, Susan. *Intellectual growth in young children.* New York: Harcourt, Brace, 1930.

Jackson, J. H. Reprint of some of Hughlings Jackson's papers on affections of speech. *Brain*, 1915, *38*, 28-190.
James, W. *The principles of psychology. Vol. I.* New York: Holt, 1890.
Jones, Mary C. A laboratory study of fear: The case of Peter. *Pedagogical Seminary*, 1924, *31*, 308-315.

Katona, G. *Organizing and memorizing: Studies in the psychology of learning and teaching.* New York: Columbia Univ. Press, 1940.
Keller, Helen *The story of my life.* New York: Doubleday, Doran, 1931.
Kendler, H. H. "What is learned?" — a theoretical blind alley. *Psychol. Rev.*, 1952, *59*, 269-277.

Klein, G. S. Need and regulation. Presented in symposium on The measurement of human motives. A.P.A., Washington, D. C., 1952.

Klein, G. S., & Schlesinger, H. Where is the perceiver in perceptual theory? *J. Pers.*, 1949, *18*, 32-47.

Koffka, K. On the structure of the unconscious. In *The unconscious: A symposium*. New York: Alfred Knopf, 1928. Pp. 43-68.

Koffka, K. *Principles of Gestalt psychology*. New York: Harcourt, Brace, 1935.

Köhler, W. *The mentality of apes*. New York: Harcourt, Brace, 1925.

Köhler, W. *The place of value in a world of facts*. New York: Liveright, 1938.

Köhler, W. *Gestalt psychology: An introduction to new concepts in modern psychology*. New York: Liveright, 1947.

Krech, D. Dynamic systems as open neurological systems. *Psychol. Rev.*, 1950, *57*, 345-361.

Krech, D., & Crutchfield, R. S. *Theory and problems of social psychology*. New York: McGraw-Hill, 1948.

Kuenne, Margaret R. Experimental investigation of the relation of language to transposition behavior in young children. *J. exp. Psychol.*, 1946, *36*, 471-490.

Langer, Susanne K. The phenomenon of language. In I. J. Lee (Ed.), *The language of wisdom and folly*. New York: Harper, 1949. Pp. 7-11.

Langer, Susanne K. *Philosophy in a new key*. Cambridge: Harvard Univer. Press, 1951. (Rev. Ed.) 1951.

Lashley, K. S. The mechanism of vision XV. *J. gen. Psychol.*, 1938, *18*, 123-193.

Lashley, K. S. The problem of cerebral organization in vision. In J. Cattell (Ed.), *Biological symposia, Vol. VII*. Lancaster, Pa.: Jacques Cattell Press, 1942. Pp. 301-322.

Lashley, K. S. Persistent problems in the evolution of mind. *Quart. Rev. Biol.*, 1949, *24*, 28-42.

Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium*. New York: Wiley, 1951. Pp. 112-146.

Lashley, K. S., & Wade, M. The Pavlovian theory of generalization. *Psychol. Rev.*, 1946, *53*, 72-87.

Leeper, R. W. The experimental psychologists as reluctant dragons. Paper read at A.P.A., Philadelphia, 1946.

Leeper, R. W. A motivational theory of emotion to replace "emotion as disorganized response." *Psychol. Rev.*, 1948, *55*, 5-21.

Leeper, R. W. Concept formation as a paradigm of learning. Paper read at Western Psychol. Assn., June, 1949.

Leeper, R. W. Cognitive processes. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 730-757.

Lewin, K. *A dynamic theory of personality*. New York: McGraw-Hill, 1935.

Lewis, Helen B. An experimental study of the role of the ego in work. I. The role of the ego in cooperative work. *J. exp. Psychol.*, 1944, *34*, 113-126.

Lewis, Helen B., & Franklin, Muriel. An experimental study of the role of the ego in work. II. The significance of task-orientation in work. *J. exp. Psychol.*, 1944, *34*, 195-215.

Lindesmith, A. R., & Strauss, A. L. *Social psychology*. New York: Dryden Press, 1949.

Littman, R. A., & Rosen, E. Molar and molecular. *Psychol. Rev.*, 1950, *57*, 58-65.

Luchins, A. S. Mechanization in problem solving. *Psychol. Monogr.*, 1942, *54*, Whole No. 248.

Luchins, A. S. An evaluation of some current criticisms of Gestalt psychological work on perception. *Psychol. Rev.*, 1951, *58*, 69-95.

Lynip, A. W. The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant. *Genet. Psychol. Monogr.*, 1951, *44*, 221-262.

McCarthy, Dorothea Language development in the preschool child. *Inst. Child Welf. Monogr.*, 1930, No. 4. Minneapolis: Univer. Minnesota Press.

MacLeod, R. B. The phenomenological approach to social psychology. *Psychol. Rev.*, 1947, *54*, 193-210.

MacLeod, R. B. Perceptual constancy and the problem of motivation. *Can. J. Psychol.*, 1949, *2*, 57-66.

MacLeod, R. B. The place of phenomenological analysis in social psychological theory. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads*. New York: Harper, 1951. Pp. 215-241.

Maier, N. R. F. Reasoning in humans. I. On direction. *J. comp. Psychol.*, 1930, *12*, 115-143.

Maier, N. R. F. Reasoning in humans. II. The solution of a problem and its appearance in consciousness. *J. comp. Psychol.*, 1931, *12*, 181-194.

Maier, N. R. F. An aspect of human reasoning. *Brit. J. Psychol.*, 1933, *24*, 144-155.

Maier, N. R. F. The effect of cortical injury on equivalence reactions in rats. *J. comp. Psychol.*, 1941, *32*, 165-189.

Maier, N. R. F. Reasoning in humans. III. The mechanisms of equivalent stimuli and reasoning. *J. exp. Psychol.*, 1945, *35*, 349-360.

Marx, M. H. The central nature of theory construction. In M. H. Marx (Ed.), *Psychological theory*. New York: Macmillan, 1951. Pp. 4-19.

Maslow, A. H. Preface to motivation theory. *Psychosom. Med.*, 1943a, *5*, (1), 85-92.

Maslow, A. H. Theory of human motivation. *Psychol. Rev.*, 1943b, *50*, 370-396.

Maslow, A. H. "Higher" and "lower" needs. *J. Psychol.*, 1948a, *25*, 433-436.

Maslow, A. H. Some theoretical consequences of basic need-gratification. *J. Pers.*, 1948b, *16*, 402-416.

Maslow, A. H. Self-actualizing people: A study of psychological health. *Personality*, 1950, Symposium No. 1, 11-34.

May, M. A. *A social psychology of war and peace.* New Haven: Yale Univer. Press, 1944.

Mensch, I. N., & Wishner, J. Asch on forming impressions of personality: Further evidence. *J. Pers.,* 1947, *16,* 188-191.

Michotte, A. The emotions regarded as functional connections. In M. I. Reymert (Ed.), *Feelings and emotions.* New York: McGraw-Hill, 1950. Pp. 114-126.

Miller, G. A. What is information measurement? *Amer. Psychologist,* 1953, *8,* 3-11.

Miller, N. E. Comments on multiple-process conceptions of learning. *Psychol. Rev.,* 1951, *58,* 375-381.

Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univer. Press, 1941.

Morris, C. *Signs, language and behavior.* New York: Prentice-Hall, 1946.

Mowrer, O. H. The law of effect and ego psychology. *Psychol. Rev.,* 1946, *53,* 321-334.

Mowrer, O. H., & Lamoreaux, R. R. Conditioning and conditionality (discrimination). *Psychol. Rev.,* 1951, *58,* 196-212.

Murphy, G. *Personality.* New York: Harper, 1947.

Novikoff, A. B. Integrative levels and biology. *Science,* 1945, *101,* 209-215.

Nuttin, J. Intimacy and shame in the dynamic structure of personality. In M. L. Reymert (Ed.), *Feelings and emotions.* New York: McGraw-Hill, 1950. Pp. 343-352.

Osgood, C. E. The nature and measurement of meaning. *Psychol. Bull.,* 1952, *49,* 197-237.

Piaget, J. Principal factors determining intellectual evolution from childhood to adult life. In *Factors determining human behavior.* Cambridge: Harvard Univer. Press, 1937. Pp. 32-48.

Piaget, J. *The psychology of intelligence.* New York: Harcourt, Brace, 1950.

Piaget, J. *Play, dreams and imitation in childhood.* Translated by C. Gattegno & F. M. Hodgson. New York: Norton, 1951.

Piaget, J. *The origins of intelligence in children.* New York: International Univer. Press, 1952.

Postman, L. The history and present status of the law of effect. *Psychol. Bull.,* 1947, *44,* 489-563.

Postman, L. Toward a general theory of cognition. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads.* New York: Harper, 1951. Pp. 242-272.

Postman, L., Bruner, J. S., & McGinnies, E. Personal values as selective factors in perception. *J abnorm. soc. Psychol.,* 1948, *43,* 142-154.

Postman, L., & Schneider, B. H. Personal values, visual recognition, and recall. *Psychol. Rev.,* 1951, *58,* 271-284.

Pratt, C. C. Operationism in psychology. *Psychol. Rev.,* 1945a, *52,* 262-269.

Pratt, C. C. Rejoinders and second thoughts. *Psychol. Rev.,* 1945b, *52,* 288-291.

Pratt, C. C. The role of past experience in visual perception. *J. Psychol.,* 1950, *30,* 85-107.

Ritchie, B. F. The circumnavigation of cognition. *Psychol. Rev.,* 1953, *60,* 216-221.

Rogers, C. *Client-centered therapy.* Boston: Houghton Mifflin, 1951.

Scheerer, M. Problems of performance analysis in the study of personality. *Ann. N. Y. Acad. Sci.,* 1946, *46,* 653-678.

Scheerer, M. Personality functioning and cognitive psychology. *J. Pers.,* 1953, 22, 1-16.

Scheerer, M. (with the collaboration of L. Immerglück and M. Buchman) An experiment in abstraction. *Confinia Neurologica,* 1949, *9,* 232-254.

Scheerer, M., Rothmann, E., & Goldstein, K. A case of "idiot savant": An experimental study of personality organization. *Psychol. Monogr.,* 1945, *58,* Whole No. 269.

Schilder, P. *The image and appearance of the human body.* London: Kegan Paul, 1935.

Schilder, P. *Goals and desires of man.* New York: Columbia Univer. Press, 1942.

Schneirla, T. C. Problems in the biopsychology of social organization. *J. abnorm. soc. Psychol.,* 1946, *41,* 385-402.

Schneirla, T. C. Levels in the psychological capacities of animals. In R. W. Sellars *et al.* (Eds.), *Philosophy for the future.* New York: Macmillan, 1949. Pp. 243-286.

Seward, J. P. The sign of a symbol: A reply to Professor Allport. *Psychol. Rev.,* 1948, *55,* 277-296.

Sherif, M. *The psychology of social norms.* New York: Harpers, 1936.

Sherif, M., & Cantril, H. *The psychology of ego involvements.* New York: Wiley, 1947.

Simpson, G. G. *The meaning of evolution.* New Haven: Yale Univer. Press, 1950.

Skinner, B. F. The operational analysis of psychological terms. *Psychol. Rev.,* 1945, *52,* 270-277.

Sloane, E. H. Reductionism. *Psychol. Rev.,* 1945, *52,* 214-223.

Snygg, D., & Combs, A. W. *Individual behavior: A new frame of reference for psychology.* New York: Harper, 1949.

Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.,* 1937, *44,* 430-444.

Spence, K. W. The basis of solution by chimpanzees of the intermediate size problem. *J. exp. Psychol.,* 1942, *31,* 257-271.

Spence, K. W. The nature of theory construction in contemporary psychology. *Psychol. Rev.,* 1944, *51,* 47-68.

Spence, K. W. The postulates and methods of "behaviorism." *Psychol. Rev.,* 1948, *55,* 67-78.

Spence, K. W. Cognitive versus stimulus-response theories of learning. *Psychol. Rev.,* 1950, *57,* 159-172.

Spence, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 690-729.

Stern, W. *Psychology of early childhood.* New York: Holt, 1930.

Stern, W. *General psychology from the personalistic standpoint.* New York: Macmillan, 1938.

Stevens, S. S. Psychology and the science of science. *Psychol. Bull.*, 1939, *36*, 221-263.

Stevens, S. S. Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 1-49.

Szekely, L. Studien zur psychologie des Denkens. *Acta Psychologica*, 1940, *5*, (1), 79-96.

Szekely, L. Knowledge and thinking. *Acta Psychologica*, 1950a, *7*, (1), 1-24.

Szekely, L. Productive processes in learning and thinking. *Acta Psychologica*, 1950b, *7*, (2-4), 388-407.

Thistlethwaite, D. A critical review of latent learning and related experiments. *Psychol. Bull.*, 1951, *48*, 97-129.

Thorndike, E. L. *The psychology of wants, interests and attitudes.* New York: Appleton-Century, 1935.

Tolman, E. C. *Purposive behavior in animals and men.* New York: Century, 1932.

Tolman, E. C. Operational behaviorism and current trends in psychology. In H. W. Hill (Ed.), *Proc. 25th Ann. Celebr. Inaug. Grad. Stud.*, Univer. S. Calif., 1936. Pp. 89-103.

Tolman, E. C. The determiners of behavior at a choice point. *Psychol. Rev.*, 1938, *45*, 1-41.

Tolman, E. C. A stimulus-expectancy need-cathexis psychology. *Science*, 1945, *101*, 160-166.

Tolman, E. C. Cognitive maps in rats and men. *Psychol. Rev.*, 1948, *55*, 189-208.

Tolman, E. C. There is more than one kind of learning. *Psychol. Rev.*, 1949, *56*, 144-155.

Tolman, E. C. A psychological model. In T. Parsons & E. A. Shils (Eds.), *Toward a general theory of action.* Cambridge: Harvard Univer. Press, 1951. Pp. 279-342.

Tolman, E. C., & Brunswik, E. The organism and the causal texture of the environment. *Psychol. Rev.*, 1935, *42*, 43-77.

Vigotsky, L. S. Thought and speech. *Psychiatry*, 1939, *2*, 29-54.

von Bertalanffy, L. The theory of open systems in physics and biology. *Science*, 1950, *111*, 23-29.

von Bertalanffy, L. Theoretical models in biology and psychology. *J. Pers.*, 1951, *20*, 24-38.

von Senden, M. *Raum- und Gestaltauffassung bei operierten Blindgeborenen vor und nach der Operation.* Leipzig: Barth, 1932.

Wallach, H. Brightness constancy and the nature of achromatic colors. *J. exp. Psychol.*, 1948, *38*, 310-324.

Wapner, S. The differential effects of cortical injury and retesting on equivalence reactions in the rat. *Psychol. Monogr.*, 1944, *57*, Whole No. 262.

Warden, C. J. *The emergence of human culture.* New York: Macmillan, 1936.

Weiss, P. *Principles of development.* New York: Holt, 1939.

Werner, H. Process and achievement — A basic problem of education and developmental psychology. *Harv. educ. Rev.*, 1937, *7*, 353-368.

Werner, H. *Comparative psychology of mental development.* New York: Harpers, 1940; 2nd Ed. Chicago: Follett, 1948.

Werner, H., & Kaplan, Edith The acquisition of word meanings: A developmental study. *Child Develpm. Monogr.*, 1950, *15*, Serial No. 1, No. 1.

Werner, H., & Wapner, S. Towards a general theory of perception. *Psychol. Rev.*, 1952, *59*, 324-338.

Wertheimer, M. *Productive thinking.* New York: Harper, 1945.

Wertheimer, Michael Hebb and Senden on the role of learning in perception. *Amer. J. Psychol.*, 1951, *64*, 133-136.

Wever, E. G., & Zener, K. E. The method of absolute judgment in psychophysics. *Psychol. Rev.*, 1928, *35*, 466-493.

White, L. A. *The science of culture: A study of man and civilization.* New York: Farrar-Straus, 1949.

Witkin, H. A. Perception of body position and of the position of the visual field. *Psychol. Monogr.*, 1949, *63*, Whole No. 302.

Witkin, H. A. *Personality through perception.* New York: Harper, 1953.

Wittreich, W. J. The Honi phenomenon: A case of selective perceptual distortion. *J. abnorm. soc. Psychol.*, 1952, *47*, 705-712.

Wood, L. Inspection and introspection. *Phil. Sci.*, 1940, *7*, 220-228.

Woodworth, R. S. Reinforcement in perception. *Amer. J. Psychol.*, 1947, *60*, 119-122.

Wright, H. F., & Barker, R. G. A method of dividing the behavior continuum. Paper read at A.A.A.S., 1951.

Young, P. T. The role of hedonic processes in the organization of behavior. *Psychol. Rev.*, 1952, *59*, 249-262.

CHAPTER 4

Psychoanalytic Theory and Its Applications in the Social Sciences

CALVIN S. HALL*
Western Reserve University

AND GARDNER LINDZEY
Harvard University

Our major intent in this chapter is first to present a précis of psychoanalytic theory and then to survey some of the applications of this theory in social psychology and the other social sciences.

We shall focus our attention upon orthodox, Freudian psychoanalysis, and mention only in passing the major revisionists. This decision to emphasize orthodox psychoanalysis does not imply a lack of appreciation for the many important contributions of other psychoanalytic theorists. It stems initially from limitations of space and beyond this from the feeling that the revisionist positions can be understood best against the background of the theory from which they evolved. Further, we are convinced that orthodox psychoanalysis, in spite of its formal shortcomings, is a far better elaborated theory than any of its close relatives, and is accompanied by a much greater wealth of pertinent empirical evidence than the revisionist positions.

Since this is primarily a chapter on theory, we have not felt obliged to review and summarize the empirical findings which in recent years have been appearing at an ever-accelerating rate from psychological laboratories and clinics. Nor shall we offer a critique of psychoanalysis, whether our own or others'. Finally, our account of Freudian personality theory is not a historical one. When we feel it necessary to discuss historical matters, e.g., Freud's early theory of anxiety, we shall do so in small type. What we shall attempt to do then is to present systematically the major constructs of Freud's general theory of personality in their final form as evolved by Freud and to discuss the relationship of these constructs to social psychology. In addition, we shall discuss the impact of psychoanalysis upon the social sciences.

PSYCHOANALYSIS AND PSYCHOLOGY

The relation between Freudian theory and contemporary psychology bears considerable resemblance to the relation between Darwin's theory of evolution and general psychology of fifty years ago. Darwin and Freud captured the imagination of thoughtful men to a point where psychologists could scarcely consider behavior uninfluenced by their views. The very act of minimizing the contributions of Darwin and Freud served to orient the deprecator toward the theory. Academic psychologists appeared restive under the influence of both of these formulators. Part of this discomfort may be directly attributed to a characteristic professional narcissism which makes it difficult to accept the contributions of persons who are not members of the profession. A further discomfiting factor was that both of these theories possessed a sweep and grandeur that evoked quick enthusiasm on the part of the interested public, while arousing dismay and aloofness in the laboratory bred investigator, with his concern for specificity and control.

In general, we might say that the reserve, or downright hostility, that academic psychologists have customarily shown to psychoanalytic theory has both a rational and an irrational component. The rational component focuses upon psychoanalysts' slight concern for carefully controlled empirical evidence, the absence of adequate empirical definitions, and the frequent statement of relationships in such a manner that they are not subject to proof or disproof. The irrational component derives from a kind of professional jealousy that strongly resents contributions made by individuals not trained in the discipline of psychology, a persistence in common-sense or normative views of behavior

---

* We wish to express our appreciation to George S. Klein and Ernst Kris for their valuable advice during the preparation of this chapter.

which are so strikingly violated by many psychoanalytic precepts, and stubborn adherence to a naïve positivism that considers nothing worth note that does not meet relatively high standards of rigor and control.

Although psychology is firmly rooted in the laboratory, while psychoanalysis is equally firmly rooted in medical practice with a tradition of observation in the therapist's chambers, they share many common antecedents. Some of these shared influences are quite general, such as the influence of Darwinism and the exposure to vigorous brands of scientific functionalism. Perhaps more important, and certainly more readily specified, is the common debt to nineteenth-century experimental physiology.

No account of the development of psychology is possible without acknowledging the great influence of Hermann von Helmholtz (Boring, 1950). Nor is any account of the development of psychoanalytic theory complete without discussion of the influence of Ernst Brücke, in whose physiological laboratory Freud spent six years prior to entering private practice. The extent of Brücke's influence upon Freud has also been extensively documented (Bernfeld, 1944; Kris, 1950; and Boring, 1950). The relation between these two ancestral figures, Helmholtz and Brücke, was not merely that of two individuals working together in the same discipline in the same era. They were close friends and had studied together under Johannes Müller at the University of Berlin. They shared a common animosity toward vitalism and an assurance that all human behavior could be accounted for in terms of physical and chemical forces. Another common antecedent was Fechner, who not only contributed in important ways to the development of psychology but also was an important influence upon psychoanalysis, as Freud's many references to Fechner testify.

We may conclude that the "new psychology" of the late nineteenth century and the first stirrings of psychoanalytic theory appeared, if not under the same auspices, at least heavily influenced by similar lines of thought. Thus, it seems fair to view the present rapprochement between academic psychology and psychoanalysis as representing the belated union of two intellectual traditions that have developed very differently but nevertheless began under importantly similar circumstances.

Yet, as we have said, one should not assume that these mutual sources of influence produced an early compatibility between Freud and the psychologists of his era. For Wundt, Titchener, Külpe, Cattell, and most other figures of this period, Freud's work was regarded as falling outside the province of psychology. Only iconoclasts such as G. Stanley Hall, James, and McDougall were able to muster any early interest in psychoanalysis and to recognize it as psychologically relevant. Even among these psychologists the interest was somewhat tentative and skeptical, as the following statement by McDougall suggests:

> In my opinion Freud has, quite unquestionably, done more for the advancement of our understanding of human nature than any other man since Aristotle . . . This judgment implies, of course, that there is much that is of value and relatively true in Freud's teaching. And yet I hold that every bit of such truth is mixed almost inextricably with error, embedded in masses of obscure implications and highly questionable and misleading propositions. (McDougall, 1936, pp. 17-18)

During the formative years of Freud and psychoanalysis, academic psychology had practically nothing to say of pertinence to the events that Freud was initially concerned with, i.e., general adaptive functioning of the human organism and the problem of mental disorder. Much of the psychology of this era consisted of speculative theorizing by individuals who were as much or more philosophers than psychologists. Such laboratory investigations as there were concerned themselves with the content of the normal, conscious, human mind. Not only did the focus upon the normal and conscious eliminate Freud and his interest in the pathological and unconscious but the methods of study employed by the laboratory investigators were equally far removed from the analysts' techniques for investigation. While the psychologist attempted to study specified aspects of behavior under controlled circumstances, where the influence of motivational and social factors could be minimized, the analyst was concerned with a very wide range of behavior (or its verbal representation) and tried to establish a setting in which personal-motivational determinants would operate at their maximum.

In spite of Freud's concern with mental disorder and the treatment process it would be a serious mistake to consider that psychoanalysis claimed to be merely an abnormal psychology. Freud made the contrary assumption:

> Psychoanalysis falls under the head of psychology; not of medical psychology in the old sense, nor of the psychology of morbid processes, but simply of psychology. (Freud, 1927b, p. 207)

This conviction is borne out in Freud's writings, where we find him perennially wrestling with the problems of normal psychological functioning. The last chapter of the *Interpretation of Dreams*, for example, is "pure" psychology in the sense that it is altogether concerned with perception, memory, motivation and thought.

With respect to social psychology, Freud did not see a sharp distinction between it and the psychology of the individual. Individual psychology, he felt, rarely could disregard the relations of the one to the many; "so from the very first, Individual Psychology is at the same time Social Psychology as well" (1921, p. 2). Consistently he felt that social psychological phenomena could be accounted for in terms of the same principles and constructs that had proven useful in representing the individual.

It is not easy to trace the path that led to eventual interpenetration of psychology and psychoanalysis. There was a mutual awareness of each other's existence quite early. Freud read widely in psychological journals and his international fame grew so rapidly during the first two decades of this century that most psychologists knew something of his theories. He was honored along with a few of his colleagues at the Clark Vigentennium in 1910 but this event was not followed by any marked increase in his academic respectability. With only rare exceptions, universities shut their doors to psychoanalysis until the thirties. The change that took place during the twenty years from 1925 until 1945 is well illustrated by examining Taylor's (1926) *Readings in Abnormal Psychology and Mental Hygiene* and Tomkin's (1944) *Contemporary Psychopathology*. These two volumes served rather similar functions in their respective eras and yet we find that the Taylor volume is virtually devoid of references to psychoanalysis, excepting those that are derisive in character, while the Tomkins volume consists predominantly of papers that bear the imprint of psychoanalysis.

One of the factors leading to increased compatibility between these traditions was the extent to which representatives of each began to deal with overlapping empirical problems. While psychoanalysts were to some extent concerning themselves with nonpathological events, the academic psychologist was moving out of the laboratory and tackling problems of socialization, adjustment, and personality. A certain amount of mutuality in empirical interest stemmed directly from World War II, where in some instances there was forced association between professional psychologists and psychoanalysts. A further outcome of the war was the development of a vigorous subdiscipline of clinical psychology which from the beginning showed a strong affinity for the concepts and some of the procedures of psychoanalysis. The growing importance of clinical psychology and its acceptance within most major universities led to a great increase in the communication between psychologists and psychoanalysts.

Independent of psychoanalysis there appeared in psychology a growing interest in motivation and motivational concepts (McDougall, Woodworth, Tolman, Allport) and this trend led to a natural reappraisal of the value of Freud's psychology of motivation. An important integrative role was played by a few individuals such as Kurt Lewin and Henry A. Murray, who from within the confines of academic institutions made earnest efforts to conduct controlled empirical research that was related to, and in part inspired by, psychoanalysis. From these efforts, psychoanalysis received a growing lustre of investigative respectability and this in turn led other, more conventional psychologists to conduct related research. During the late thirties the efforts that were made to bring about a rapprochement between Hull's reinforcement theory and aspects of psychoanalysis (Dollard, Miller, Mowrer, Sears) brought many experimentalists into at least temporary contact with Freud's conceptions of behavior. Finally, the period of the thirties was a time of unrest. Radical ideas were welcomed, social change was a matter of interest and appreciation, axioms were being challenged. In this setting the controversial nature of Freudian theory was more closely attuned to the spirit of the times than before.

For all these reasons, psychoanalysis has become an area of great interest for psychologists. Today it seems evident that Freud's position as a major figure in the development of psychology is secure. The benefits of this influence may still be debated, but its existence is no longer open to question.

ENERGY, LIBIDO, AND CATHEXIS

ENERGY

The concept of energy is no less difficult to explicate in psychology than it is in the other sciences. It is doubly hard to get hold of Freud's notion of psychical energy because in his early theories he adopted a position from which he appears to have receded in subsequent reformulations. The concept of libido, its definition as sexual energy, and its protean character in mental life are well known to every student of psychoanalysis. Yet the assumption that Freud emphasized the sexual impulse to the relative exclusion of other impulses, which is justified by Freud's pre-1920 writings, does not hold for his final theoretical scheme. Freud did not help to clarify the changed state of affairs because he tended to keep old concepts while developing new ones. Thus when he formulated the polar instincts, life versus death, he was reluctant to part with a single form of energy, libido, which had worked so well for the polarity of the sexual instincts vs. the ego instincts. One of the reasons Freud abandoned the latter polarity was that inasmuch as libido governed both groups of instincts, they were not really opponents but partners in the enterprise of life. As a result of grouping the sexual and ego instincts under the life instincts, libido seemed to lose its close tie with the purely sexual instincts, to become desexualized, and to function as a general psychical energy which could run all of the psychological processes, and not just those that are in the service of reproduction. In any event, libido is identified by Freud as the motive force of eros or the life instincts.

What then is the energy of the death instincts? With this question Freud wrestles again and again. In one such encounter, he tries to preserve a monistic libido theory by deriving the energy of the destructive instinct from libido. He hypothesizes that there is a displaceable energy which can be used indifferently for attaining erotic or destructive aims, but that this energy is itself desexualized libido (Freud, 1923a). However, neither this solution nor any other one satisfied Freud, because in his final summing up he states "there is no term analogous to 'libido' for describing the energy of the destructive instinct" (Freud, 1940, p. 22). This statement suggests that Freud came to believe that there are two forms of energy, one which governs the life instincts and all the derivatives therefrom and one which governs the death instincts and all the derivatives therefrom. It is not clear from Freud's writings whether these two forms are different or the same in quality. He *does* say that "nearly all the energy with which the mental apparatus is charged comes from the inborn instincts" (Freud, 1920, p. 5), but this statement does not help us to decide whether he believed in one kind of energy or as many kinds as there are instinctual sources. In one of his writings Freud *does* imply a belief in a general instinctive energy because he criticizes Jung for making libido synonymous with psychic energy in general (Freud 1905, p. 6). In this connection, Freud says that the differentiation of the libido from the other forms of instinctive energy represents one of the great achievements of psychoanalysis. These statements suggest that Freud did postulate a monistic psychical energy. But we cannot be sure. We turn now to a short account of libido theory, which will be elaborated in later sections of this chapter.

LIBIDO

Freud's clearest statement of what he meant by libido is found in *Three Contributions to the Theory of Sex* (Freud, 1905). In this monograph he defines libido "as a force of variable quantity by which processes and transformations in the spheres of sexual excitement can be measured" (p. 611) and distinguishes it from general psychic energy by virtue of its origin in the erogenous zones of the body. Since any part or organ of the body can be an erogenous zone, i.e., seat of sexual excitement, the sources of libido are multifarious. The erogenous zones, Freud hypothesizes, are sensitized by chemical substances (hormones) secreted by the gonads.

Libido either remains stored up in the ego or it is invested in sexual objects. In its stored-up state it is called *ego libido,* and in its invested state it is called *object libido.* When the investment is withdrawn, the libido may take the ego as an object and it is then known as *narcissistic libido.* There is a reciprocity between ego libido and object libido, so that the more there is of one the less there is of the other.

CATHEXIS

In one of his writings, Freud characterizes psychoanalysis as "a dynamic conception which reduces mental life to the interplay of reciprocally urging and checking forces" (Freud, 1910b, p. 107). These forces are *cathexes* and *anticathexes.* Because cathexis refers to the way in which energy is distributed throughout the personality it is a central concept in psychoanalytic theory. Yet its meaning is not always clearly understood.

This much is plain, cathexis refers to an investment or charge of energy. But what is it that becomes invested or charged with energy? Freud uses the term in such a way as to lead the reader to believe that either external objects or the ego itself can become invested with energy. If it is the former it is called an *object cathexis,* if it is the latter, an *ego cathexis* or *narcissism.* Narcissism results from the process of *identification* (incorporation) by which the ego assimilates the object and secures for itself thereby the cathexis which formerly resided in the object. Actually this is a very loose and misleading description of the cathexis process, and should not be taken in any literal sense. One does not invest energy in an object nor does one normally incorporate an object, except for food, liquids, and the ingredients of air. Rather, the investments are made in psychological processes and their products. When Freud speaks of cathexes and anticathexes he is referring to the distribution of energy throughout the whole psychic apparatus. That is to say, he views the person as a complex energy field and he seeks to understand the person in terms of the way his energy is distributed and for what purposes and with what consequences. In the most general sense, psychic energy may be said to be used for the purposes of living and dying, but this tells us no more than that a person lives and dies.

Let us begin at the beginning. At first there is only a primitive psychic apparatus which is operated by a small quantity of mobile energy generated by the instincts. This energy dissipates itself in smooth muscle and striped muscle activity before it has a chance to accumulate. At this stage, one cannot speak of cathexis, since the energy is not invested in anything. It is spent as soon as it is earned. Environmental circumstances in conjunction with maturational developments force this reflex-like apparatus to undergo modifications in its structure, with the result that energy is used to operate such psychological processes as perception, memory, and thought. At this stage one can speak of the formation of cathexes because energy is invested in psychological processes and is not spent immediately in purely muscular and visceral activities. Energy flows into the perceptual, mnemic, and ideational systems and the contents of these systems, percepts, memories, and ideas, respectively, become charged with energy. A high charge of energy is called a hypercathexis. As long as this charge remains on a content of a psychological system it is said to be bound. When the gateway into the motoric is opened, the energy becomes mobile and discharges itself through activity (Freud, 1900).

From this account, it can be seen how misleading it is to speak of an object cathexis. Properly speaking, one can cathect a perception, or a memory, or a conception of an object but not the object itself. Similarly, one can cathect a perception or a memory or a conception of one's self but not the ego itself. Therefore, when Freud speaks of object cathexis and ego cathexis or object choice and narcissism he means the activation of a perception, memory, or thought whose reference is an external object and the activation of a perception, memory or thought whose reference is the self. Identification is simply a transformation of a cathexis for the perception of an object into a cathexis for a perception of the self. Incorporation, by the same token, represents a shift in cathexis from a psychological content which has an outer referent to a psychological content which has an inner referent. The objective is transformed into the subjective.

An anticathexis is a checking force in mental life. Energy is used for the purpose of resisting or inhibiting the discharge of energy through the motoric system. Since the aim of the primitive personality is to expel energy as soon as it is generated, any postponement of discharge requires an investment of energy in blocking or delaying processes. An example of a blocking process is repression, of a delaying process is thinking.

Freud distinguishes between two kinds of *frustration,* one external, the other internal (Freud, 1915a). In external frustration the goal object is not available for reasons beyond the person's control. This type of frustration does not involve anticathexis and might better be called privation or deprivation. Internal frustration proceeds from within the personality structure and consists of inner inhibiting forces, i.e., anticathexes. The attempts of the personality to overcome or circumvent these internal frustrations constitute the major task of psychological functioning. Freud observes that internal frustration does not come into existence until external frustration has prepared the ground for it, by which he means that the person has to experience external privations and deprivations before he can develop inner controls. This is the basis for the formation of the ego and superego.

Frustration, whether of the internal or external variety, places a strain upon the personality because the damming up of energy causes an increase in tension. The frustrated person can retain his mental stability by taking direct action on the world and wresting satisfaction from it, or he can use the energy for ends which do not have anticathexes attached to them (Freud, 1912).

From this brief account of cathexis and anticathexis we can see that these concepts are the very essence of Freud's dynamic personality theory, and that the measurement of these forces constitutes the chief task of the psychoanalytic investigation.

THE SUBDIVISIONS OF PERSONALITY

In Freud's final formulation of psychoanalytic theory (Freud, 1940), the primary provinces of personality are identified as the *id*, the *ego*, and the *superego*. The original distinction between conscious and unconscious which were the theoretical pillars of psychoanalysis prior to the formulation of the theory of the id, ego, and superego by Freud, first made public at the Berlin Psychoanalytic Congress in 1922, no longer enjoy this distinction. In fact, Freud relegates consciousness and the two kinds of unconscious, preconscious and unconscious proper, to the status of qualities. Thus, it is necessary in reading Freud's publications prior to 1922 to substitute "id" for "unconscious" and "ego" for "preconscious" and "conscious" in many places. In the present account of psychoanalytic theory we have made these substitutions whenever it seemed appropriate to do so. Although the contents and processes of the id are exclusively unconscious, one is not warranted in treating id and unconscious as equivalents, in spite of the widespread convention of doing so. The contents and processes of the id have the quality of being unconscious but the id is *not* the unconscious, except in a historical sense. By identifying that agency of the mental apparatus which Freud once called the unconscious as the id, we believe that much of the confusion which occurs when one reads pre-1922 Freud and post-1922 Freud should disappear. Moreover, although the entire id has the quality of being unconscious, much of the ego and a greater part of the superego have the same quality. Thus it is to compound an error to use "unconscious" as a synonym for the id.

Although Freud finally demoted the concept of the unconscious to a mere quality of mental life (Freud, 1933), the importance of the concept in the history of psychoanalytic theory, quite apart from its later identification as the id, should not be minimized. Freud felt that in order to establish psychology as a natural science, he had to destroy the popular notion held by psychologists and philosophers of the nineteenth century that mental life consisted solely of that which is conscious. A psychology of consciousness limited itself, so Freud thought, to a very thin slice of mental life. Consciousness is only one quality or attribute of that which is mental, and a highly unstable one at that. Although later Freud was to characterize unconsciousness and preconsciousness in the same terms, except that they were more stable qualities than that of consciousness, by focusing upon the unconscious as the larger part of what is mental he aspired to establish psychology as a complete science of the mind (Freud, 1938).

Since Freud does not equate mental with consciousness, his use of the term "mental" is unorthodox. The mental consists of somatic processes, which are largely unconscious, and they are similar in kind to all the other natural processes of which science has obtained knowledge (Freud, 1940). When Freud writes that "our scientific work in psychology will consist in translating unconscious processes into conscious ones, and thus filling the gaps in conscious perceptions" (Freud, 1938), he means that the goal of psychology is to discover the essential nature of unconscious processes and their contents and to make this knowledge available to people so that they can become aware of that which they cannot experience directly. It is hardly necessary, one might add, to have a science of direct experience, since the goal of any science, whether it be physics or psychology, is to make known that which cannot be known directly.

THE ID

The id is the obscure, inaccessible, and unorganized part of personality. What little is known about it has been learned from study of the dream work and from investigations of the formation of neurotic symptoms. It has no direct relations with the external world but is in direct contact with somatic processes. Freud calls the id the "true psychic reality" (Freud, 1900), because it represents only the inner world and is not contaminated by objective (external) reality.

Origin of the id. The id is the oldest of the mental systems. It consists of everything that is inherited, that is fixed in the constitution (including the instincts), and that is present at birth. Experiences that are repeated often and intensely in individuals of successive generations are deposited in the id and become a part of the inheritance of succeeding generations. This is what Freud means when he says that the vestiges of countless egos are stored in the id (Freud, 1923a). In other words, Freud subscribes to the doctrine of acquired characters as the basis for the formation of the id. Throughout Freud's writings one finds references to the signal role of phylogeny in laying the foundation of the human mind.

Contents of the id. "Instinctual cathexes seeking discharge — that in our view is all that the id contains" (Freud, 1933, p. 105). This terse statement requires amplification. For Freud, an instinct is a distinctly psychological concept and not a biological one. It is defined as the mental (ideational) representation of a somatic (visceral) process (Freud, 1905). These visceral processes are called "instinctual impulses," and it is a looseness of terminology, Freud believes, to confuse instinctual impulses with the instinctual ideas which arise from them (Freud, 1915a). The id contains ideas of immediate gratification, i.e., *wishes*. All of the energy of the id is invested in wishes, so that the id can only do one thing, namely, to wish. Thus *wishes* can be substituted for *instinctual cathexes seeking discharge* and the statement would then read "the id contains only wishes." It is essential to bear in mind the distinction between impulse and wish. Impulse is the somatic counterpart of a wish and *pari passu* wish is the psychological counterpart of an impulse.

The id knows no reality other than a subjective, autistic reality. It is without values, ethics, and morality. The contents of the id are immortal, for they do not alter with the passage of time. Nothing in the id is past or forgotten. The laws of logic and reason do not hold for the id, so that contradictory wishes can exist side by side without negation of one by the other. The energy of the id is in a more fluid state than is the energy in the other regions of the personality (Freud, 1933). Freud's metaphor for the id is "a cauldron of seething excitement."

Functions of the id. The sole purpose of the id is to provide for the free and untrammeled discharge of quantities of excitation which arise from somatic sources. In fulfilling this sovereign purpose the id is said to obey the *pleasure principle* and to make use of the *primary process*.

The pleasure principle is a tendency whose function is to render the psychic apparatus as a whole free from any excitation (this is called a state of Nirvana), or in lieu of this ultimate achievement to keep the amount of excitation constant and as low as possible. The aim of the pleasure principle differs in no respect from the universal tendency of all living matter, which is to strive to return to the quiescence of the inorganic world (the death instinct). Although the pleasure principle, through its agent, the primary process, tries to counteract excitations whose sources lie in the external world by discharging them as fast as they arise, its main task is to deal with the *inner* increases in tension by keeping them at a minimum through providing a direct path from excitation to discharge (Freud, 1920).

While the id knows no other principle of functioning than that of pleasure, the other agencies of personality, viz., the ego and superego, also bend to the will of the pleasure principle. The ego and superego may modify the pleasure principle in various ways but they cannot nullify it. Freud considered it to be a question of the greatest theoretical importance how it is ever possible for the pleasure principle to be overcome. "The consideration that the pleasure principle requires a reduction, or perhaps ultimately the extinction, of the tension of the instinctual needs leads to problems that are still unexamined in the relations between the pleasure principle and the two primal forces, Eros and the death instinct" (Freud, 1940, p. 109).

The psychic process employed by the id is called the *primary process* (Freud, 1900). This process consists of discharging tension as soon as any tension reaches the id. The first and most primitive operation of the primary process is to expend energy immediately in motor activity. Examples of this are the emptying of the bladder's contents when the pressure on the sphincter reaches a certain level, or the mass activity of a baby when it has colic.

For several reasons, this reflex action phase of the primary process does not serve effectively the dictates of the pleasure principle. In the first place, many of the somatic sources of tension do not have built-in tension-reducing mechanisms, as do urination and defecation. Most of them are like hunger, the tension of which can be reduced only by a specific kind of commerce (eating) with a specific kind of object (food). Moreover, a reflexly discharging organism is bound to encounter many rebuffs from the external world, so that instead of finding the pleasure which it seeks it finds pain, i.e., increased tension.

Consequently, the primary process develops another method for discharging tension. This method consists of producing a hallucination of the desired object. Thus the hungry individual attempts to rid himself of tension by cathecting a hallucinatory perception of food.

This function of the id presupposes that the organism has had previous experiences with particular goal objects for satisfying particular needs and that it is capable of perceiving and remembering. It must be able to perceive food and to remember that it has perceived food, so that when it becomes hungry again the tension of hunger can activate the memory of food and reinstate the associated perception (hallucination) of food. This is what Freud means when he says the primary process strives for discharge of excitation by establishing an identity of perception. Identity of perception refers to the equivalence of a centrally aroused image of food with a perception of food aroused by an actual food stimulus.

The process which begins from a state of tension (need) and proceeds to the cathexis of perception by way of memory is what Freud calls wish-fulfillment. A wish may be defined as a cathexis for the perception of an object that will satisfy the wish, i.e., discharge the tension. Accordingly, the sole psychological function with which the id is endowed is that of generating wishes.

Wishing is said to be regressive in character, since it reverses the usual course of the excitation from perception to memory. Instead, the excitation flows backward from mnemic system into the perceptual system. The resulting hallucination proves, in the long run, to be no more satisfying than was the earlier reflex discharge of tension through the motoric. Consequently, the primary process is superseded by another one which Freud calls the *secondary process*. The development of the secondary process goes hand in hand with the formation of a new structure in the personality, the ego. The secondary process is the process by which the ego accomplishes its main objective, to wit, to be ruled by the reality principle instead of the pleasure principle. We shall take up the discussion of this process in the next section.

THE EGO

The ego is the organized subdivision of personality. It is the executive, embracing all of those functions which are thought of as being distinctly psychological, viz., perception, learning, memory, and reasoning. It occupies a central position in the personality, standing between the id and the external world, the superego and the external world, and the id and the superego. Like the administrative head of a democratic political system, it both governs the other agencies of personality and is governed by them. Moreover, it has the responsibility of conducting all transactions with the outer world.

Origin of the ego. Under the influence of the external world acting through the perceptual system one part of the id undergoes a special development in order to act as an intermediary between the id and the external world (Freud, 1940). This is the ego. Heredity influences the ego by way of the id. It may be said that even before the ego comes into existence its subsequent lines of development, its predispositions, and its reactions are already determined by heredity (Freud, 1937a). Thus, the ego is not the product of experience alone but of experience acting upon a prepared organism.

The boundary between the ego and the id consists of resistances or anticathexes. In a healthy person, the amount of energy which goes into maintaining the ego-id barrier is kept to a minimum, while in a neurotic person a great deal of energy is utilized for this purpose.

Functions of the ego. The primary function of the ego is to protect the life of the individual against the dangers which arise in the external world. To fulfill this objective, a number of differentiated psychological processes and systems are developed in the ego. It becomes aware of the external world by means of stimuli acting upon the perceptual system. It stores up experiences of stimuli in the mnemic system. It learns either to adapt to the external world or to bring about changes in the external world through activity. Freud believed that mastery of the environment constitutes the highest achievement of the ego (Freud, 1926a). Thinking and judging, along with perception and memory, serve to inhibit the immediate discharge through motor pathways of excitations reaching the ego. The ego, therefore, is said to control the path of access to motility.

The ego's aim of self-preservation requires that it gain control over the demands of the instincts in the id, since the immediate gratification of the instincts may endanger the life of the person. The ego decides whether an instinct shall be allowed to obtain satisfaction. If the decision is in the affirmative, the time and circumstances for its satisfaction are determined by the ego. If the decision is in the negative, the ego has the task of suppressing the instinctual excitations or of diverting them into other channels.

The ego binds the mobile energy of the id by transforming it into a quiescent, tonic charge (Freud, 1920). This is accomplished by the development of a *secondary process* which takes the place of the primary process of the id. The secondary process arranges mental activities in sequential order and tests their correspondence with reality. This is called *reality testing* (Freud, 1923a). The secondary process may even cathect that which induces pain if by so doing it is in accordance with the reality principle by which the ego is governed. After the ego has completed its work of thought, it removes the inhibition and permits the release of energy through motility.

In the last analysis, the ego in a normally well-adjusted person is the managing director of the personality. It synthesizes the forces acting upon it from the id, from the superego, and from the external world into a harmonious unity under the wise and benevolent jurisdiction of the reality principle.

Reality principle. As a consequence of the failure of the primary process to provide gratification, there is formed out of the id a new mental agency, the ego, which is governed by a new principle, the *reality principle* (Freud, 1911). The aim of the reality principle is to inhibit the discharge of excitation until the appropriate object in or condition of the environment for the satisfaction of the need is discovered. The substitution of the reality principle for the pleasure principle does not mean that pleasure is forsaken. The reality principle serves the pleasure principle by postponing an immediate gratification, which is likely to bring pain in its wake, for a later gratification in which the risk of suffering pain is considerably reduced. All of the complicated mental activity which is enjoined by the reality principle represents a roundabout way to the ultimate goal of pleasurable gratification.

The inauguration of the reality principle calls for a succession of adaptations in the mental machinery, which are required by the important role that the external world plays in the satisfaction of needs. The perceptual system develops more precise and accurate ways of reflecting external reality. A special faculty, attention, is instituted whose function it is to search the outer world for items of knowledge which can be used in the interest of need satisfaction. There is introduced a system of notation, language, whose task is to preserve the results of attending and perceiving in the mnemic system. Judgments are made according to the criteria of truth and falsity instead of those of pleasure and pain. Motor discharge is restrained by means of the process of thought, which is an experimental way of acting that involves smaller quantities of cathexis and less expenditure of them. There is even provided for within the ego a mode of thought activity free from the demands of reality testing which consists of creating fantasies and daydreams. This is a self-conscious and rather more controlled use of the primary process, which the ego recognizes for what it is, namely, a pleasurable pastime, and does not confuse with reality. Finally, there is a change in the motoric system so that it becomes more capable of manipulating and modifying reality through skillful and efficient action.

Secondary process and reality testing. The primary process of the id, it will be recalled, consists of discharging energy according to the model of a reflex apparatus. Input and output of excitation are immediate reflex actions. If the output cannot be effected because the object which will end the inner tension is absent, then the primary process hallucinates the desired object. When this operation proves to be ungratifying an ego process is developed. The excitation is diverted by this process into the activities of thinking and judging, which seek to find the desired object in the real world. The process of reasoning out a way of discovering the appropriate goal object is called the *secondary process* (Freud, 1900). When one thinks he has found a satisfactory solution to a problem, the verification of the solution by trying it out is called *reality testing*. The secondary process is mental problem solving or experimental action, as Freud labels it, whereas reality testing is the behavioral means for verifying the truth of a thought or perception, i.e., its existence in reality.

Perception and memory. All psychic activities proceed from stimuli acting upon sense organs and terminate in innervations of motor organs. That is, the psychic apparatus has a sensory end and a motor end. At the sensory end lies the perceptual system which receives excitations from the external and internal sense organs. The function of the perceptual system is to distinguish among sensory qualities. The sensory qualities of perceptions which originate in the external world are many and varied but those which come from the viscera are restricted to the pleasure-pain series, i.e., feelings (Freud, 1900).

Perception is not merely a passive process. Small amounts of energy are sent out periodically from the reservoir in the ego into the perceptual system and sample the stimuli acting upon the sense organs (Freud, 1925a). The functioning of the perceptual system requires energy, as does every system of the personality. When the energy is withdrawn the perceptual system ceases to function.

The relation of perception to consciousness and the question as to whether there are unconscious perceptions are not clearly formulated by Freud. In some of his writings he seems to equate perception with consciousness, so that everything which is perceived must have the quality of consciousness, while in other writings Freud implies that there are unconscious perceptions. If one accepts the hypothesis that the quality of consciousness is produced by a hypercathexis of psychic material, then it is quite conceivable that some perceptions might have insufficient cathexis to raise them to consciousness yet have enough to permit their existence unconsciously. Faced with the necessity of making a decision between these two alternatives, we have decided in favor of the view which ascribes to all perceptions the quality of consciousness. Our decision rests upon the following statement made by Freud, ". . . as soon as the cathexis is withdrawn [from the perceptual system], consciousness is extinguished and the functioning of the perceptual system comes to a standstill" (Freud, 1925b, p. 180).

One may perceive, i.e., be conscious of, thoughts as well as sensory excitations. This is accomplished by the association of ideas via the function of speech with memory traces of the perceptions of written or spoken words. Since ideas can become conscious, it is necessary to have a way of distinguishing between a perception whose source is a thought and a perception whose source is a sensory stimulation. The former is what Freud terms subjective reality, the latter objective or material reality. In order not to confound the two, the ego resorts to *reality testing*. This process consists of performing some action upon the external world. If the perception is made to disappear by such action, then the origin of the perception is known to lie in the external world. If such activity makes no difference to the perception, then the perception originates within the person and it does not have objective reality (Freud, 1916).

The perceptual system retains nothing; it has no memory. Accordingly, another system, the mnemic system, which lies next to the perceptual system in the sequence of psychic activities from sensory input to motor output, is postulated. This system consists of memory traces which are the permanent deposits of perceptions.

Memory traces may be associated with one another. Freud conceives of association as a lessening of resistance and a smoothing of the way from one trace or *mem-element* to another trace or mem-element. In terms of cathexis, association consists of an unobstructed flow of energy from one trace to another trace. Associations are formed when the percepts which left the traces occur together (law of contiguity) or when there is a similarity between the percepts (law of resemblance) or when other laws of association operate on the percepts. Freud thought that there might be many mnemic systems or *mem systems*, as he dubbed them, each of which would be founded upon a particular law of association (Freud, 1900).

Memories that are painful to the person acquire a charge of anticathexis which renders their entry into consciousness difficult or impossible. They are said to be *repressed*. A memory which is not painful in itself may be repressed because it is associated with another memory which is painful. Thus we forget neutral or innocuous experiences because they are tied by cathexis to disagreeable experiences. This is the essential feature of Freud's theory of forgetting (Freud, 1904).

In a brief article published in 1925 Freud draws a parallel between the operation of a child's toy, the "mystic writing pad," and the functioning of perceptual-mnemetic systems (Freud, 1925b). This toy consists of an outer celluloid cover on which the child writes with a stylus, a thin sheet of waxed paper on which the impression made by the stylus is recorded, and a slab of wax or resin which is soft enough to permit indentions to be made in it, so that an impression may be made on the covering waxed paper. When the child wishes to erase what he has written he merely lifts the outer cover and the thin undersheet, and the writing disappears. Consequently, the pad can be used over and over again.

The parallel between this toy and the perceptual and mnemetic systems is as follows. The outer cover represents the perceptual system's protective screen against stimuli, the waxed paper is like the perceptual system, which receives perceptions but does not retain them, and the under thickness of wax is analogous to the mnemetic system, which preserves the permanent traces of the perceptions. By imagining the toy to be written on and erased in rapid alternation, one has a picture of how perception operates.

Freud believed that memory traces could be inherited, so that the mind preserves not only what happens to it in the course of an individual's life but also what happened to the ancestors of the individual. Freud felt that by the analysis of dreams one might obtain knowledge of the archaic memory of man (Freud, 1900).

Thought and judgment. The major objective of the ego is to inhibit the outflow of excitations in the interest of the reality principle. The postponement of discharge is accomplished in large measure by the ego's capacity for thought. Thought is to be regarded as an experimental action which involves only a small expenditure of energy (Freud, 1925a). It makes use of experiences stored in memory. Thinking employs the secondary process for securing an identity of thought with reality.

When thoughts are congruent with reality, the ego opens the sluice gates of the motoric system and an appropriate response is made.

Incorrect thinking, such as occurs in condensations, displacements, confusions, and errors, results from an intrusion of id processes. The ideas of the id are concrete and thing-bound. Thus a person who confuses a hole with a vagina does so because the id reacts solely to the physical similarity of the two objects (1915b). The symbolism of dreams is but one example of the id's influence on thought. Autistic thinking in general, except insofar as it stems from the special fantasy-making region of the ego, is id-dominated thinking (Freud, 1900).

Judgment has two functions. It may assert or deny that a thing has a particular property or it may affirm or negate that a particular image exists in reality. The original property of an object which comes under the scrutiny of judgment is whether the object is likely to cause pleasure or to cause pain. The prototype for judging pleasure and pain is that of affirmation through incorporation, e.g., swallowing, or negation through expulsion, e.g., spitting out. When something is affirmed as being pleasurable the love instinct is in control, whereas when something is expelled as being painful the destructive impulse is ascendant. With the development of the ego, qualities of objects other than the pleasure-pain ones assume importance for the proper functioning of the ego.

The highest development of judgment consists of determining whether a psychic state is real or not. It accomplishes this aim by rediscovering in the world, through perception, the referent of the thought. If it cannot be found in the world the thought is a delusion and is discarded. It is precisely in this way that a scientist attempts to test his hypotheses, i.e., his thoughts about some aspect of reality (Freud, 1925a).

The ego in relation to the id. The ego, as has already been noted, originates out of and is the organized portion of the id. Consequently, the relationship between them is a very intimate one. There is no natural antagonism between the ego and the id, and in the mature, well-integrated person there is really no division between them (Freud, 1926a). They work as a unit under the guidance of the reality principle, which seeks to gratify the wishes of the id in a manner that does not violate the dictates of reality. If a split does occur between the ego and the id, the weakness of the ego then becomes apparent. When this happens the ego is forced to substitute rationalization for logical thought to gloss over the conflicts between the id and reality, and to engage in other pretenses (Freud, 1933). But when the ego retains firm ties with the id and is indistinguishable from it, then the ego displays its great power (Freud, 1926b).

The ego has no energy of its own but must borrow it from the id. Sometimes it resorts to trickery to withdraw further amounts of energy from the id. One such trick is that of identification. Since all object cathexes proceed from the id, the ego, by identifying itself with an object, recommends itself to the id in place of the object and seeks thereby to attract the energy of the id onto itself. Consequently, the ego takes into itself a large number of precipitates of former object cathexes (Freud, 1933).

The ego's relations with the id are influenced by the superego, which is partly a reaction formation against the id. Consequently, the ego is forced into the awkward position of serving two masters who are at odds with each other.

THE SUPEREGO

The third major subdivision of personality, the *superego,* represents morality. It is the last of the provinces of personality to be formed and reflects the standards of the society in which the child is raised. The superego strives for perfection.

The superego has two aspects, one the *conscience,* the other the *ego ideal.* The ego ideal sets the standards of ethical conduct, while the conscience acts in the capacity of a judge who metes out punishments for infractions of the rules.

Origin of the superego. The superego is formed out of the ego in childhood as a consequence of the dissolution of the oedipus complex (Freud, 1940). This dissolution is accomplished by substituting identification with the parents for feelings of love and hate toward the parents. The parents or, more precisely, their qualities as punishing and rewarding figures in the early life of the child, are introjected into the ego and become a separate institution, the superego, which performs the same functions as the parents once did. The punishing qualities of the parents become the conscience and the rewarding qualities become the ego ideal. When the ego does not live up to the expectations of the ego ideal, conscience punishes the ego by making it feel guilty. When the ego does manage to fulfill the ideals of the superego, it is rewarded by feelings of pride. The superego is modelled upon the earliest parental images, when the parents were thought to be perfect and omnipotent. These early images of the parents affect the child's later relationships with other authority figures and with the laws and ideals of society (Freud, 1933).

The superego is derived primarily from auditory impressions received from the parents in the form of commands, accusations, threats, and exhortations (Freud, 1923a). As a consequence of its origin in auditory perceptions, the superego remains accessible to consciousness, although it is for the most part unconscious. A person can remember, in substance at least, what his parents once said to him because it has been incorporated and is permanently retained by the superego. The voice of conscience is a transcription of the parents' voices. However, the cathectic energy of these verbal remnants does not originate from the auditory perceptions as such but from sources in the id. That is, since the superego represents an introjection of the earliest object choices of the id, i.e., the parents, the cathectic energy in the superego is derived from the strength of the cathexis for the parents. In other words, the severity of the superego is determined not by the severity with which the parents treated the child but by the intensity of the cathexis for the parents and the consequent strength of the resistance against the dissolution of the oedipus complex. The stronger the oedipus complex is, the stronger is the superego, since the energy of the former flows into the latter when the oedipus complex is abandoned (Freud, 1940).

In general, parents follow their own superegos in bringing up children. The result is that the child's superego is really not a reflection of the parents' behavior but is a reflection of the parents' superego. In this manner, the traditional values and ideals of a society are handed down from generation to generation (Freud, 1933).

Functions of the superego. The functions of the superego are far simpler than those of the ego. Representing as it does parental authority, the superego behaves toward the ego as the parents once did toward the child. When the ego breaks a superego rule, the superego punishes the ego by making the person feel guilty in the same way that a child is made to feel guilty by parental admonition. The superego withdraws its approval as the parents once withdrew their love, so that the reproaches of conscience correspond to the child's fear of losing his parents' love. The superego commands, corrects, and judges the ego in a wholly parental fashion. In a similar way, the superego bestows its favor upon the ego when the ego defers to the superego. The ego is filled with pride when it measures up to the ego ideal, just as a child feels proud when his parents praise his virtuous conduct. The conscience has the specific task of keeping the ego under surveillance to see whether it lives up to the ego ideal (Freud, 1940).

All in all, the chief function of the superego is to limit instinctual gratifications, not in accordance with the reality principle which guides the ego's control over the impulses, but in accordance with perfectionistic and unrealizable standards imposed by the parents or, more correctly, by the superegos of the parents. This aim of limiting instinctual gratification is due to the circumstance that parental authority is exercised mainly in suppressing or regulating sexual and aggressive impulses. Parental rules of conduct are categorical imperatives and, therefore, it is the categorical imperative which rules in the superego.

The superego in relation to the id. The id and the superego have one thing in common; they both represent influences of the past on the individual. In the case of the id the past is contained in the individual's biological inheritance, while in the case of the superego the past is represented by the social tradition in which the individual grows up. Since Freud believed that social tradition can become a part of one's biological inheritance, the past of the id and the past of the superego are considered to have much in common. This belief moved him to observe that "what belongs to the lowest depths of the mind is changed through formation of the ego ideal (superego) into what we value as the highest in the human soul" (Freud, 1923a, p. 48).

Leaving to one side these biological speculations, the superego can be said to be the receptacle for the earliest object choices of the id. Thus the id cathects the superego as it once did the introjected objects of the superego and uses it as an ally against the ego. The superego can act as the representative of the id in relation to the ego (Freud, 1923a).

However, the superego contains more than the object choices of the id. It also contains reaction formations against these choices. The superego not only says in effect "you ought to be like your father," but it also says "you may not do all that he does," e.g., have sexual relations with the mother (Freud, 1923a).

The relations of the ego and superego. Although the superego develops out of the ego as a consequence of the resolution of the oedipus complex, its relation to the ego is a burdensome one for the ego. In the first place, the superego develops during a stage when the ego is still so weak that it is unable to cope realistically and effectively with the demands of this new province upon it. In the second place, since the superego contains the oldest object choices of the id it is a fellow conspirator of the id against the ego. And thirdly, introjected materials of the superego were once real persons in the environment of the individual, so that the superego may be said to be a representative of the outer world. In every sense, the superego stands in opposition to the ego. When the ego is strong it can perform its synthetic function and effect a fusion of the two forces, the id and the world, which the superego represents, but when it is weak, as it is when the superego is formed, it is more likely to be ruled by than to rule the superego. By establishing the superego the ego masters the oedipus complex, but in so doing it places itself again in subjection to the id and to a small extent to the external world as well. Because the ego is the representative of the internal subjective world, conflicts between the ego and the superego arise over the distinction between what is real and what is mental (Freud, 1924).

The expression of tension between the ego and the superego manifests itself in feelings of guilt, whereas the expression of harmony between the two is felt as pride. The ego is constantly measuring itself against the perfectionistic standards of the superego and striving to attain them. In a sense, the displacement of libido to the ideals of the superego and the attainment of these ideals represents a recovery of the primary narcissism which was lost when the ego was formed (Freud, 1914). The ego can now love itself by living up to its ideals, a kind of bootstrapping which is not at all uncommon in the relations between the various agencies of personality. On the other hand, the ego can gain the condemnation of the superego by identifying with objects and persons who have been rejected as being unworthy by the superego. Thus self-love and self-hate both have their sources in the kinds of identifications made by the ego.

As long as all goes well in the life of a person, the superego is lenient and lets the ego do pretty much as it pleases, just as a parent, when he is in good spirits, takes an indulgent attitude toward the child's conduct. However, should a calamity occur, the superego goes into action, holds an inquisition, discovers the sin, heightens the standards, imposes abstinences, and demands penance from the ego (Freud, 1930). This sudden transformation of the superego reminds one of a similar transformation in a parent who discovers that his child has broken a favorite possession. The superego pursues its own ends and is independent of the ego as regards the energy at its disposal.

In concluding this section on the subdivisions of personality it is well to bear in mind what Freud said about the dividing up of personality into separate provinces. He felt that there are no sharp boundaries between the three major provinces; rather they are like areas of color which shade off into one another. After separating them "we must allow what has been separated to merge again" (Freud, 1933, p. 110). Freud conceived of the total personality of the normal person as working as a unit and not as three separate systems in opposition to one another. The latter state of affairs characterizes the neurotic individual who is out of joint with himself. For such a person, psychotherapy consists of effecting a psychosynthesis.

The Mental Qualities

A content of the mind may have one of three qualities. It may be *conscious, preconscious,* or *unconscious.* Only those contents which are the result of ego processes can be conscious. Neither the id nor the superego can make anything conscious. On the other hand, not all that the ego produces is conscious. Most of it, in fact, is preconscious.

The contents of the ego which may become conscious are perceptions, feelings, and thoughts (ideas). A perception is a response to a particular quality of an object in the external world. When a person says that he sees a red table, i.e., is conscious of, aware of, or has an image of a red table, he is perceiving a stimulus property, red. Feelings are qualities of the pleasure-pain series which have their sources in the processes of the id.

Freud was not certain as to just what processes in the id were responsible for feelings of pleasure and pain. At one time he felt that an increase of instinctual tension was painful, a decrease pleasurable, but he soon saw that this could not be true, since there are some tension increases, notably those which occur in sexual excitement, which are clearly pleasurable. Subsequently he decided that pleasure and pain depend upon a qualitative factor in the tension, e.g., the rhythm of tension changes, rather than upon a simple quantitative lowering or raising of tension (Freud, 1924).

A feeling can never be preconscious or unconscious, although the processes which give rise to it are preconscious or unconscious (Freud, 1923a). Thoughts which are originally preconscious become conscious by being connected with the memory traces of speech symbols.

These memory traces have been etched in the mind by the perception of verbal and written material. In order for a thought to become conscious it must have a connecting link with language (Freud, 1939). Infantile memories cannot become conscious because they have no associations with language symbols.

The quality of consciousness is the only one of the three qualities known to us directly. What it means for a mental content or process to be preconscious or unconscious is a "shrouded secret" (Freud, 1940). Freud believed, however, that it has primarily to do with the charge of energy (cathexis) on the percept or idea. If the cathexis is weak the idea is preconscious, and if the cathexis is still weaker the idea is unconscious. However, one must take into consideration not only the cathexis of the process itself but also the possibility of there being another process whose cathexis acts in opposition to the first one. The net cathexis that remains after the second one (the anticathexis) has been subtracted from the first one will determine with what quality, whether conscious, preconscious, or unconscious, the process is endowed.

In Freud's earlier theorizing, when unconscious, preconscious, and conscious were conceptualized as regions of the mind, he pictured the anticathexes or resistances as functioning at the boundary between the conscious and preconscious, and the boundary between the preconscious and unconscious (Freud, 1900). These resistances constitute the censorship operating upon material striving to cross the border in the direction of becoming conscious. The terminology used for describing this earlier model of the mind perseverates in Freud's later writings and causes some confusion. However, it seems quite clear to us that when Freud changed his model from a conscious-preconscious-unconscious one to an id-ego-superego one, he intended to do away with the conception of boundaries or barriers between consciousness and unconsciousness, and to replace a geographical plan with a dynamic one. The conscious or unconscious quality of a percept or idea is determined by the sum of positive and negative charges of energy. Consequently, one may speak of an idea as being nearer or farther away from consciousness, depending upon its charge and opposing charges. An idea that has a strong unconscious quality, i.e., "is buried deep in the unconscious" is either one which has a weak charge on it or one which has a strong anticathexis, or both. The qualitative categories of unconscious and preconscious are merely convenient classes for grouping material which differs quantitatively but not qualitatively. In other words, there is a continuum from the weakest possible unconscious idea to the strongest possible conscious one. Where one draws the lines between the three qualities is purely arbitrary.

THE INSTINCTS

In this section we consider Freud's motivational theory which, like his theory of the structure of personality and his theory of anxiety, underwent such drastic transformation in the early twenties that few of Freud's associates were willing to accept the reformulation (Freud, 1920). Psychoanalysts had become used to the older libido theory, so that the subordination of sex to the hegemony of the life force and the introduction of a powerful death force upset the theoretical apple cart. Freud was not able to put the apples back in the cart again to his own satisfaction let alone that of his colleagues. To this day, the relationship of the new motivational theory to other psychoanalytic theories has not been satisfactorily worked out. The death instinct, in particular, is an embolism in the psychoanalytic body which has yet to be absorbed. It may be fairly said that the resistance to the acceptance of a death instinct outweighs the resistance to any other feature of psychoanalysis, even that of infantile sexuality, which aroused so much animosity in the early years of psychoanalysis.

DEFINITION AND ORIGIN OF INSTINCT

An *instinct* is the mental representative in the id of an inner somatic source of stimulation (Freud, 1905). A stimulus arising from an exciting process in an organ or bodily tissue is called a *need.*

Freud tends to confuse the terms *need* and *instinct* in his writings, although it is clear to us that he intended to distinguish them as separate concepts. Need is a biological concept, instinct a psychological one. Need is to instinct as stimulus is to perception.

The study of needs or the sources of instincts lies outside the province of psychology. Instinct is "a borderland concept between the mental and the physical, being both the mental representative of the stimuli emanating from within the organism and penetrating to the mind, and at the same time a measure of the demand made upon the energy of the latter in consequence of its connection with the body" (Freud, 1915c, p. 64).

A need differs from an external stimulus in several ways. The source of the two is different, since a need is located within the body while an external stimulus is located in the outer world. A need operates as a more or less constant force, whereas a stimulus usually has a momentary impact. The most important difference is that one can escape from an external stimulus by flight, whereas one cannot escape from a need by this means. External stimuli are more easily mastered than are the instinctual stimuli and the latter make greater demands upon the person because they require more complicated activities in order to be satisfied. Freud believed that needs and not external stimuli are the more important cause of the development of the nervous system. However, he asserts that there is nothing to prevent one from assuming that the needs themselves are phylogenetic residues of external stimulation (Freud, 1915c).

CHARACTERISTICS OF INSTINCTS

In addition to their sources in bodily need states, instincts have an impetus, aim and an object. The impetus of an instinct is its motor element, i.e., the quality of impulsion or activity which characterizes all instincts. The aim of an instinct is to abolish the tension created by the need. Thus an instinct is said to be conservative in nature, since it seeks to re-establish a basal tension level which existed before the need arose, i.e., a return to an earlier state. In accomplishing this aim the individual experiences pleasure. The object of an instinct is, as a rule, a particular kind of commerce with an object in the external world. The aim and the object of an instinct can be altered or exchanged for other aims and objects. However, the object is the most variable thing about an instinct and substitutions of objects are more easily made than substitution of aims. In general, then, an instinct can be thought of as a sum of energy which imparts direction to psychological processes (Freud, 1933).

Freud says that instincts are conservative by nature. By this he means that they seek to repeat that which has gone before (*the repetition compulsion*) or if development has taken place they strive to return to an earlier stage of development (*regression*).

NUMBER OF INSTINCTS

It is possible to distinguish an indeterminate number of instincts but they all belong to or are derived from two basic groups, the *life instincts* and the *death instincts.* The instincts within each group are distinguished from one another by virtue of the fact that each has a distinctive somatic source, e.g., hunger vs. sex, and a distinctive special object, e.g., eating vs. copulation.

The dichotomy of *ego instincts* and *sexual instincts* was Freud's earliest instinct theory. The ego instincts had as their aim the self-preservation of the individual, while the sexual instincts aimed at preservation of the species through reproduction. As early as 1914 Freud expressed dissatisfaction with this theory but he was not inclined then to give it up entirely (Freud, 1914). Instead, he elaborated the concepts of *ego libido* and *object libido,* which in a sense constituted a new instinct theory and paved the way for the final formulation of life instincts and death instincts (Freud, 1923b). Freud recognized that the ego was the main reservoir of libido, having taken this function away from the id, and that it invested its energy in object cathexes. However, a goodly portion of it always remained in the ego and was said to be used narcissistically because the ego cathected itself. Moreover, there was a constant transformation of ego libido into object libido and of object libido back into ego libido. This fact made Freud suspicious of the basic nature of the distinction between the two kinds of cathexes, object versus ego, with the consequence that he banished his original instinct theory, and created a new one. The ego instincts were fused with the sexual instincts to give the life instincts, and a new group of instincts, the death instincts, which bear no relation to any early features of the theory, were postulated (Freud, 1920).

EROS OR THE LIFE INSTINCTS

The life instincts have as their goals the preservation of life and the preservation of the species. They accomplish these aims by establishing ever greater unities both in connection with the synthetic function of the ego and in the formation of libidinal relations with objects. The person becomes more highly integrated within himself and with others (Freud, 1940).

Since the principal feature of an instinct is that it seeks to return to an earlier state, one wonders what the earlier state is which Eros tries to reinstate. It would have to be a state of unity which subsequently had been split up and to which the individual wished to return. Freud was not satisfied with this analysis because he felt that there is nothing in the actual history of living substances to support such a hypothesis, other than the supposition that the birth of life out of inorganic matter was accomplished by the shattering of an inorganic unity. If this were so, the life instincts and the death instincts would have the same aim, viz., to return to the inorganic, and there would be no need to differentiate between them (Freud, 1940).

The best known and most minutely analyzed group of life instincts are the sexual instincts. The psychoanalytic literature is replete with studies of the sexual instincts, so much so that for many people the distinguishing and ubiquitous feature of psychoanalysis is its pansexualism. Prior to 1920 this may have been true; after that date pansexualism became a misnomer.

The general characteristics of the sexual instincts are as follows (Freud, 1915c). They are numerous, arising from many bodily sources. At first each sexual instinct acts more or less independently of the other sexual instincts but as the individual matures they achieve a more or less complete synthesis in the service of propagation. In the early stages of development they serve the aims of the ego instincts (self-preservation) from which they only gradually detach themselves and acquire their own aims. In their choice of objects, also, they at first adopt the objects of the ego instincts.

The sexual instincts are very plastic. They can change their aims and their objects and substitute for one another, so that a particular sexual instinct can be vicariously satisfied by another sexual instinct (Freud, 1933).

The sexual instincts do not appear suddenly at puberty but go through a course of development beginning in earliest infancy (Freud, 1905). The erogenous zones of the body, which come into prominence one at a time, are the key to this development. An erogenous zone is a portion of the skin or mucous membrane from which unpleasant feelings of sensitiveness arise. The removal of these feelings by manipulation of the zone, e.g., sucking or massage, evokes pleasure.

The first such zone is the mouth, from which pleasure is derived independently of nourishment. Oral pleasure is obtained in two ways, by stimulation of the lips and oral cavity and by using the teeth for achieving aggressive aims. Oral erotic pleasure develops before oral aggressive pleasure, because the latter has to await the appearance of teeth before it can be fully satisfying, although even before the teeth appear, aggression finds an outlet through spitting-out behavior. The two oral stages are often referred to as oral incorporation and oral sadism.

The second erogenous zone to come into prominence is the anus, which does so by virtue of its relation to the excretory function. The anus has two modes of functioning, retention and expulsion, either of which yields pleasure. To a far greater extent than is true of the oral period, pleasure derived from the anal modes depends upon the discharge of destructive energy. Toilet training procedures, parental discipline in the form of reprobations and approbations, and external controls over bodily pleasures enhance the aggressiveness of the child during this stage.

The third region of the body from which the child obtains pleasure through tactual stimulation, i.e., masturbation, is the genital zone. Since the male genitals are of primary importance in this stage, this is called the phallic stage of libido development.

These three stages, *oral, anal,* and *phallic,* which are not to be thought of as appearing and disappearing in a one-two-three order but which overlap and occur simultaneously, constitute the pregenital development of the child during the first five years of life. At the end of this period, the child goes into latency, from which he emerges at puberty and pursues a final libidinal development. This final development consists of the organization of all of the sexual functions in the service of reproduction. Some of the earlier cathexes are restrained or repressed, some are integrated into the sexual act as forepleasure, while others are transformed into character traits or displace their aims through the process of sublimation. The principal distinction between the pregenital and genital sexual functions is that the former are autoerotic and narcissistic while the latter are alloerotic and object oriented.

Freud points out that much of our highly valued cultural heritage has been acquired at the cost of inhibiting and repressing the sexual instincts. The energy which is used for sexual gratification is compelled to find new channels. Love is one such channel. All manifestations of love, whether they be in the form of self-love, love between parents and children, friendship and love for humanity, or devotion to objects and abstract ideas, are expressions of the same basic instinct, sexual love (Freud, 1921). These manifestations of love are usually diverted or prevented from attaining the culmination of sexual union, and in fact their very strength and tenacity are dependent upon the inhibition of this aim, since the attainment of sexual union dissipates libido and weakens emotional ties.

THE DEATH INSTINCTS

"The goal of all life is death" (Freud, 1920, p. 47). By this, Freud means that there is resident in living organisms a compulsion to return to the inorganic state out of which living matter was formed. The aim of the death instincts is just the opposite of eros. Whereas eros forms ever larger unities, the death instincts work to destroy unity.

The death instincts do not put themselves on display in behavior as do the life instincts. For the most part they continue their inner task of bringing about the death of a living thing inconspicuously (Freud, 1923a). However, aggression turned outward or aggression turned inward, which are derivatives of the death instincts, are far from being inconspicuous in behavior. The derivative instinct of destruction diverts the energy of the death instincts away from the person's own dissolution and projects it out on the world. When the superego takes form, some of the aggression is used by the superego against the ego in the form of self-destruction.

INTERACTION BETWEEN LIFE INSTINCTS AND DEATH INSTINCTS

Rarely is any act the product of a single instinctual impulse. Rather, the rule is that the two basic instincts of life and death, and their innumerable derivatives, either work together in some sort of synthesis or they work against each other and reveal a conflict of interests. Any action is, therefore, the product of a number of motives (Freud, 1932).

The death instincts are rendered relatively harmless by entering a coalescence with eros. The libido of eros neutralizes the self-destructive energy by diverting it outward in the form of mastery, domination, aggression, and the will to power. Love may use mastery in gaining possession of an object. On the other hand, love may acquire a sadistic quality by being combined with the destructive instinct. There are innumerable ways in which instincts can and do combine and the unravelling of the motives which produce a given action is often a difficult and always a complicated business. But this unravelling, if it is pursued to completion, always brings the investigator to the basic life and death instincts (Freud, 1924).

When the life and death instincts are pursuing their respective aims more or less independently, that is, one at a time, then the gratification of the one makes it possible for the other to put in an appearance. When eros has been satisfied, the death instincts are likely to come to the fore, and vice versa. This accounts for the observation that love is often succeeded by hate, and hate by love.

THE EGO IN RELATION TO THE INSTINCTS

Although Freud never clearly formulated the relations of the instincts to the ego and the superego, one gains the impression from a careful reading of his various thoughts on this subject that the ego is the agent of the life instincts and their derivatives, and the superego is the agent of the death instincts and their derivatives. This distinction is by no means a strict one, and it is much clearer in the case of the superego than it is in the case of the ego.

The primary aim of the ego, it will be recalled, is that of maintaining life, and all of its vast and complicated elaborations of psychic processes is directed toward this superordinate goal. When Freud says that life is a roundabout way to death, he must mean that the roundaboutness is provided by the development of an ego. For an id without an ego would soon destroy itself.

The ego's relation to the sex instincts is one of trying to find satisfaction for them in line with the reality principle. This involves finding appropriate objects in the external world, of incorporating these objects by the mechanism of identification, of sublimating (desexualizing) the sexual instincts, and of utilizing numerous other techniques for dealing with the sexual instincts. There is always a danger that the control may go too far in opposing the purpose of eros, with the consequence that the death instincts may seize the opportunity to assert themselves when the sexual instincts are inhibited. If this is the case, then the ego becomes indirectly the agent of death (Freud, 1920).

However, the ego also possesses the means of dealing with the primitive death wish of the id. It can project it outward in the form of aggression or sublimate it in the form of mastery over the world or diminish its force by turning it into masochism, which is a form of suffering without dying (Freud, 1924).

THE SUPEREGO IN RELATION TO THE INSTINCTS

The superego, based as it is upon external authority, has closer ties with the death instincts than with the life instincts. Perhaps the most characteristic feature of the superego is the aggression which it displays against the ego. Although the model for this aggression may lie in the punitive parent, the real source of it lies in the death instinct of the id. From this source to its ultimate appearance in the superego, the death wish of the id undergoes considerable modification. Through the action of the ego on it, by the mechanism of projection, the wish for one's own death is transformed into aggression toward others. The external authority figure, by imposing prohibitions on the manifestations of the life force, especially that of sex, becomes a particularly fitting target for aggression. This defiance elicits a counterattack from the parent, e.g., castration threats and loss of love, so that the child in the interest of his own safety inhibits his aggression and develops a superego which assumes the role of the punitive parent toward the ego. The superego may seek to destroy the ego through harsh and cruel treatment, so that the death wish, by this extensive detour, may finally achieve its purpose. Depression, melancholy, and suicide are manifestations of the death wish which has found a new base for its operations in the superego. The ego becomes masochistic under the influence of a sadistic superego (Freud, 1924).

We can now understand why it is said that the superego is corruptible. It is corruptible in the sense that a primitive id wish for death obtains gratification under the guise of the superego's morality. Moreover, the superego's sadistic morality can be projected outward in the form of cruel treatment of others, which is rationalized as being justified by a high moral purpose.

ANXIETY

The three basic emotional states which Freud discusses are *love, anger,* and *fear* or *anxiety.* Love and anger each have their own energy sources in the life and death instincts, respectively. Anxiety, on the other hand, is derived from the same circumstances of life which instigate the formation of the ego and superego. It is a by-product, albeit an extremely important one, of psychological development and, in turn, anxiety causes further transformations in personality. The avoidance of anxiety becomes a very important motive in the shaping of character and in the institution of defensive measures.

DEFINITION AND FUNCTIONS OF ANXIETY

Anxiety is a painful affective state which is differentiated from other states of unpleasure, e.g., tension, pain, and melancholy, by a specific quality which is contributed by automatically innervated responses of the heart, lungs, and other internal organs through the perceptual system to consciousness. As so defined, anxiety and fear are interchangeable terms. Freud seems to have accepted the James-Lange theory of emotions, although nowhere does he say so.

The sole function of anxiety consists of its being a danger signal to the ego, so that the ego may institute measures to terminate the danger (Freud, 1926b).

TYPES OF ANXIETY

There are three types of anxiety, *reality* or *objective* anxiety, *neurotic* anxiety, and *moral* anxiety. These types do not differ from one another in any qualitative way; they are all experienced by the ego as a painful emotional state. They differ in respect to their sources. Reality anxiety results from the ego's relation to the external world, neurotic anxiety from the ego's relations to the id, and moral anxiety from the ego's relation to the superego (Freud, 1926b, 1933).

Objective or reality anxiety. Objective anxiety is an emotional reaction to a perception of danger in the external world. A danger is defined as an external condition which one remembers as having had a traumatic effect in the past. By *traumatic* is meant something which paralyzes the person so that he is not able to make a coping response to the source of the intense, excessive excitation.

When the anxiety reaction occurs it can take one of two directions. Either the person avoids the danger by fleeing from it or by mastering it, or he does nothing and the anxiety mounts until it overwhelms the ego and the person regresses to a state of infantile helplessness.

Objective anxiety appears first in development and arises out of the condition of helplessness of the baby. Any condition which portends a return of this helplessness, e.g., threat of castration or loss of love, will initiate the anxiety reaction. The external danger must be perceived as one which is related to a situation of helplessness that has been experienced earlier in life (Freud, 1926b, 1933).

Neurotic anxiety. Neurotic anxiety is an affective reaction to a perception of danger in the internal world. Specifically, it is a fear of the potential harm that may be done by the unleashing of the instincts. Neurotic anxiety is a derivative of reality anxiety, since fear of the id is basically a fear of what the id may cause the person to do.

Neurotic anxiety can be observed in three forms. There is a free-floating, general apprehensiveness which is ready to attach itself to any more or less suitable environmental circumstance. Neurotic anxiety also manifests itself in phobic reactions. A phobia is characterized by an intensity of anxiety which is not warranted by the objective situation. A person who becomes paralyzed with fright when he is on a high place, even though he is fully protected from falling, is obviously reacting to something more than the hazard of an objective situation. He is reacting to an unconscious desire to jump, the effect of which is to throw him into a state of horror at the thought that the wish might be realized. Fear of high places, or any other phobia for that matter, may also be an expression of moral anxiety. A person is afraid that his superego will take advantage of an external situation in order to punish the ego. Finally, there are anxiety states which seem to have no justification whatsoever in any external danger. Inexplicable panic reactions are of this nature.

Neurotic anxiety is often projected outward as objective anxiety because it is easier to protect oneself from an external danger than from an internal danger. Obviously, this projection offers no permanent solution, since the person continues to carry the danger around with him (Freud, 1933).

Moral anxiety. Moral anxiety is an emotional reaction to a perception of danger from the superego. The ego considers itself to be in danger when the superego is angry with it or threatens to punish it or ceases to love it. Moral anxiety expresses itself as feelings of shame and guilt.

Since the ties between superego development and the oedipus and castration complexes are so intimate it is not surprising that moral anxiety also has close ties with neurotic and reality anxiety. The superego, on the one hand, abhors the instinctual impulses in spite of the fact that these same impulses may exploit the superego for their own ulterior purposes, and on the other hand it is the introjected voice of parental authority which, at one time, actually produced objective anxiety by threats and punishments. Thus moral anxiety gathers force by being a synthesis of neurotic and objective anxiety (Freud, 1926b).

ORIGIN OF ANXIETY

Anxiety is produced primarily by an accumulation of excitation resulting from external stimulation which is so excessive that the ego is not able to bring it under control, i.e., to bind it. The event which is the prototype for this source of anxiety is birth, when the young organism is likely to be overwhelmed by an excessive amount of excitation.

Anxiety may also be produced by perception of an impending traumatic situation. The anxiety, in this case, acts as a signal to the individual that he is in danger. The origin of the signalling function of anxiety is one that lies in the individual's past experience with traumatic situations. He learns to anticipate trauma and to do something about it by paying attention to the incipient anxiety which the perception of danger initiates. By developing apprehension, the ego is in a position to ward off the danger and avoid the accumulation of intolerable amounts of excitation.

Objective or reality anxiety appears developmentally before either neurotic or moral anxiety. Freud observes that in each stage of development there is a particular condition for fear which is peculiar to that stage. For example, in the oral dependent phase, the great fears are the loss of the love object or the loss of love, in the phallic phase it is fear of castration for boys, and during latency the child fears his conscience. Freud believes that fear of castration plays the greatest role in the anxiety of boys, and that fear of loss of love is the decisive factor in the anxiety of girls.

Neurotic anxiety develops in one of two ways. As a result of inhibition, excessive tension may accumulate in the id until it exerts such a strong pressure on the ego that the ego is reduced to a state of helplessness and reacts to its helpless condition with anxiety. Or the ego may experience neurotic anxiety when an instinctual demand entails a real external danger. The person is afraid of an impulse in the sense that he is afraid of the trouble the impulse might get him into should it be expressed. An instinctual demand of moderate strength is not dangerous in itself; it becomes so when it is likely to lead to an objective danger.

Moral anxiety arises whenever the ego perceives that it has failed to live up to the standards set by the superego. Since the superego is the representative within the personality of external authority figures, the ultimate source of moral anxiety is an external danger (Freud, 1926b, 1933). In fact, the ultimate source of all anxiety is the perception of an external danger.

Until 1926 Freud believed that anxiety was a consequence of the damming up of energy by repression. In his essay, *Inhibitions, Symptoms, and Anxiety* (1926b), he reversed himself and announced that anxiety was the cause of repression rather than its effect. In a later lecture he wrote "To our surprise we discovered that it is not the repression that creates the anxiety but the anxiety is there first and creates the repression" (Freud, 1933, p. 120). Freud said that the division of the personality into id, ego, and superego forced him to adopt this new position with regard to the problem of anxiety.

After reversing his earlier theory, Freud continued to assert that repression made available energy which might be converted into anxiety, although he felt that it was of no importance which portion of the general energy was employed for this purpose (Freud, 1926b).

THE EGO IN RELATION TO ANXIETY

The ego is both the seat of anxiety and that part of the personality which attempts to abate the anxiety. The way in which the ego deals with anxiety is as follows. When an external stimulus acts upon the primitive mental apparatus and is perceived as being painful, restless motor activity takes place until the painful perception disappears. When the percept reappears, the same thing happens, so that a trace of the experience is laid down in the mnemic system. Eventually, the beginning of a painful excitation arouses apprehension or anxiety, which is a signal to the ego that more and stronger pain is in the offing unless steps are taken to avoid the painful situation. If the ego feels helpless, the anxiety mounts until the ego either rises to the occasion in some effective way or surrenders to the anxiety, i.e., has an anxiety attack or becomes completely immobilized (Freud, 1900).

Through the processes of learning and retention, the amount of anxiety which is necessary to act as a signal is reduced to a minimum. The efficiency of the ego consists, in part, of anticipating dangers on the basis of reduced cues. Anxiety may also be dealt with by defensive measures instituted by the ego.

THE SUPEREGO IN RELATION TO ANXIETY

The superego is responsible for the development of moral anxiety in the ego. When the ego is conscious of the sources of this type of anxiety it is experienced as guilt feelings. Consciousness of guilt arises before there is a conscience and consists of fear of external authority. When the superego originates, the fear of external authority is converted into fear of the superego's authority and criticizing function. The ego feels guilty when it perceives that it has failed to act in accordance with the wishes of the superego (Freud, 1930).

A mere intention to do something that transgresses the superego's code is sufficient to produce feelings of guilt. Feeling guilty before any misdeed has occurred is a signal to the ego to obliterate the intention in order to avoid more intense condemnation from the conscience. The stronger the superego is in relation to the ego, the more the person is predisposed to experience guilt. This follows from the fact that a strong superego demands more instinctual renunciation and every time the ego feels any pressure for gratification from the id, the ego becomes flooded with guilt. It follows that the person with the most fully developed conscience will experience the greatest amount of moral anxiety.

Defenses of the Ego

In its normal functioning the ego makes use of its various subsystems to avert danger when the anxiety alarm is sounded. It attends to and tries to apprehend the nature of the danger; it thinks over the various possible courses of action; it draws upon memory to recall former ways of handling the danger; it considers, judges, and evaluates; finally it takes some appropriate action. These normal functions operate more or less successfully in dealing with external threats, but they are not as well suited for protecting the ego against internal dangers, i.e., id and superego demands. Accordingly, the ego initiates additional measures whose purpose is to deal with conflicts between the id and the ego, and the superego and the ego. These measures are called the *defense mechanisms* of the ego. Freud characterized them as being "pathological," since they accomplish their mission by falsifying and distorting reality and by using energy which could be employed to better advantage by the rational systems of the ego (Freud, 1937b).

Freud first used the term "defense" in 1894 (Freud, 1894) but soon abandoned it in favor of the term "repression." In 1926 he revived the earlier concept. At that time he defined defense as a general name for all of the techniques, other than the normal ones, which the ego uses in putting down conflicts. The term "repression" was reserved for a special method of defense. Thus, in reading the writings of Freud which appeared between *The Aetiology of Hysteria* (1896) and *Inhibitions, Symptoms, and Anxiety* (1926b) it should be borne in mind that the term "repression" is sometimes used in a restricted sense, i.e., as merely one of many defense mechanisms, and is sometimes used in a general sense, i.e., as a designation for all of the defense mechanisms.

REPRESSION

Perhaps the earliest defense initiated by the ego against the id is that of repression, which frustrates the discharge of instinctual energy by setting up a block, i.e., an anticathexis, so that the instinct cannot become conscious or directly expressed in overt behavior. That is, repression acts directly upon the aim of the instinct. No other defense operates upon the aim, so that repression is more sharply differentiated from the other defenses than they are from one another.

Freud differentiates two phases of repression.

The first phase, which he calls *primal repression*, is a denial of entry into consciousness of the ideational presentation of the instinct. That which has always been in the id stays there. However, the id has the capacity for transforming its content into forms which are capable of being smuggled past the resistances into the ego. When the ego unmasks the disguised material and discovers an instinct presentation, it automatically institutes the mechanism of repression. This repression of the derivatives of material which is subject to primal repression is called *repression proper*. Repression can also be used against material which, although not harmful in itself, has come into association with material that is judged to be harmful. Freud points out that a storehouse of primal repressions in the id, which at no time have been in the ego, is a prerequisite for repression proper (Freud, 1915d).

Repression is established and maintained by anticathexes, generated by the ego and used to resist the driving forces of the id. Anticathexis is the sole mechanism used in primal repression. In the case of repression proper, there is, in addition to resisting forces, a withdrawal of cathexis from the instinctual idea which changes a preconscious quality into an unconscious one. Freud suggests that the cathexis which is withdrawn from the idea is the very one which is used in generating an anticathexis (Freud, 1915b).

The mechanism of repression is carried out by the ego under the influence of anxiety. The ego becomes aware of an incipient instinctual demand which, if permitted to discharge itself, would endanger the safety of the individual. The cathexis of this demand must be removed. If the ego is strong, the cathexis can be removed by assimilating it into the organization of the ego. If the ego is weak, painful feelings are generated and repression is automatically brought into play. By repressing the instinctual demand, the painful feelings are eliminated. The earliest and most basic repressions arise out of traumatic situations, when the ego is faced with excessive pressures from the id (Freud, 1933).

What happens in the id during repression? According to an earlier hypothesis, Freud thought that the energy of the repressed impulse was converted directly into anxiety. Upon his repudiation of this theory of anxiety, Freud re-examined the problem of what happens in the id when repression occurs and concluded that the effects might be varied. The repressed impulse might retain its libidinal cathexis and continue to exist unchanged in the id. Or the libidinal cathexis might be entirely removed from the impulse and diverted into other channels. The impulse could then be said to have been destroyed in the id. This is what happens normally to the oedipus complex. Or there might be a regression of the libidinal organization to an earlier stage of development. In any event, by outlawing material from the ego, it becomes subject solely to the laws which govern the id, especially the principle of the repetition compulsion, so that the ego very seldom succeeds in undoing its repressions (Freud, 1933).

The ego, although it often institutes repression upon orders from the superego, may turn repression against the superego itself and secure oblivion for a sense of guilt.

Repressed material returns to the ego and becomes conscious or expresses itself in behavior, e.g., symptoms or errors, when (1) the strength of the anticathexis is diminished or (2) the strength of the id cathexis is increased (Freud, 1939). There are many conditions, such as illness and sleep, which diminish the power of the resistances, and there are many conditions, such as energy changes at puberty and the climacteric or the addition of energy by some circumstance of life which makes contact with the repressed material, that will increase the strength of the cathexis. Also, because the id is capable of working upon its material by the processes of distortion, displacement, condensation, and symbolization, it can transform it so that the ego is not aware of its source. When the ego does become aware of the danger inherent in the material, the ego breaks out with anxiety and repression is again instituted. It is possible, of course, for the repressed material to express itself in the form of more or less permanent symptoms and character traits which may never be recognized for what they are by the ego. A sort of secondary anxiety may arise from the presence of a disabling symptom or trait, but this secondary anxiety is always more welcome than the anxiety which would result were the individual to discover the real origin of the symptom or trait. A trait or symptom gains satisfaction for a repressed impulse without incurring anxiety. That is why they are so hard to dislodge once they have taken hold in the person (A. Freud, 1936).

Another way that the subject matter of a repressed idea can enter consciousness is for the ego to deny its acceptance of the idea. Whenever a person uses a negative, e.g., "I don't want to hurt you," he has actually had to assert the positive before it could be interdicted. The example should be read in two parts. "I want to hurt you" represents the repressed idea, and "No, I don't" is the denial (Freud, 1925a).

REGRESSION

In *The Interpretation of Dreams* Freud distinguishes three kinds of regression. One he calls *topical,* by which is meant the substitution of the primary process for the secondary process. This is the kind of regression that is responsible for dreaming. Another kind of regression is called *temporal,* which refers to the appearance of older psychic formations in place of more recent ones, and a third is denoted as being *formal* because it describes the substitution of primitive modes of expression and representation for the customary and normal ones. Having made these distinctions, Freud then proceeds to say that they are all basically one, "for that which is older in point of time is at the same time formally primitive and, in the psychic topography, nearer to the perception-end" (Freud, 1900, p. 497).

Regression qualifies as a method of defense because it is employed by the ego to satisfy the demands of the id in such a way, i.e., autistically, that the ego is not really endangered. There is a substitution of fantasy for reality or, putting it in another way, a symbol is created for an anxiety-producing referent.

REACTION FORMATION AND PROJECTION

These two favorite mechanisms are discussed together because one of them, reaction formation, involves a reversal of a conscious affective state while the object of the instinct remains the same, and the other, projection, involves a reversal of subject and object while the affective state remains the same. In reaction formation, the anxiety-producing feeling, "I hate you," is changed into the anxiety-reducing feeling, "I love you." Subject (I) and object (you) do not alter their form, but the conscious feeling of love is now just the opposite of hate. Of course, the hate still exists unconsciously, i.e., it is repressed, but its place in consciousness has been taken by a completely fraudulent feeling. In projection, "I hate you" becomes "You hate me." In this case the true feeling of the instinct is expressed openly, but the change in object, i.e., you into me, constitutes the defensive distortion.

It is easy to see why projection is such ? prevalent defense when one remembers that, on the whole, the ego is better prepared by experience and practice to deal with external dangers. By fabricating a "real" enemy, the person feels more adequate to take command of the situation. He can, for example, destroy or attempt to destroy his enemy, which brings satisfaction for his aggressive impulse without incurring a feeling of shame. For this reason, projection is probably the most effective of the defense mechanisms. By contrast with it, reaction formation is much less satisfying, since the person has to pretend a feeling he does not have and the expression of which gives him no gratification for the real instinct. Still it does act as a bulwark against anxiety.

SUBLIMATION

Whether or not sublimation should be classed as an ego defense is a perennial question.

Anna Freud lists sublimation as an ego defense, but she says it "pertains rather to the study of the normal than to that of neurosis" (A. Freud, 1936, p. 47). S. Freud does not make a definite pronouncement as to the status of sublimation, although he does observe in *Civilization and Its Discontents* (1930) that sublimation is one method available to man for guarding against pain.

By sublimation, Freud meant that sexual energy (libido) is diverted from its sexual objective to higher and finer cultural goals. Culture is, in fact, acquired at the cost of the sexual instincts. Law and order, for example, constitute sublimations of anal eroticism. Freud says that many people have the capacity to direct part of their sexual powers into their work. "The sexual impulse is particularly suited to yield such contributions because it is endowed with the capacity of sublimation, i.e., it has the power to exchange its nearest aim for others of higher value which are not sexual" (Freud, 1910a, p. 46). Freud admits that he does not know what it means to say that an aim is higher or finer, and that he is using these terms only in a metaphorical sense (Freud, 1930).

OTHER DEFENSES

In addition to the ones already discussed, Anna Freud (1936) discusses five other defenses. They are isolation, undoing, introjection, turning against the self, and reversal. For a discussion of these defenses the reader is referred to Anna Freud (1936) and Fenichel (1945).

CHARACTER FORMATION

Freud's well-known formula for the formation of character states that the permanent character traits are "either unchanged perpetuations of the original impulses, sublimations of them, or reaction formations against them" (Freud, 1908a, p. 50). To this he later added the notion that character is also created by incorporation of early parental functions, and identifications with parents at a later date and with other persons in authority. Thus a boy will assume masculine character traits by identifying with his father, while a girl will acquire feminine character traits by identifying with her mother. A child models himself after the same-sex parent in an attempt to achieve the same kind of love relationship with the opposite-sex parent that the same-sex parent has. He also strives for identification in order to allay his anxiety arising from open aggression toward and fear of the same-sex parent. This is known as identification with the aggressor and plays a major role in character formation. Identifications are also made with people with whom one has had object relationships which have been abandoned (Freud, 1933).

The prototypes of what later become character traits occur during the pregenital phases of development. A person whose character is heavily saturated with perpetuations, sublimations, reaction formations and identifications of the oral phase is said to have an oral character. In like manner, if his character is formed around a nucleus of anal prototypes, he is said to have an anal character. However, the character of most people will include precipitates from all of the early phases of libidinal development.

Freud makes a sharp distinction between character and neurotic symptoms, although not between character and defenses. In character formation repression either does not operate at all or it easily reaches its objective, which is to replace the repressed impulses with reaction formations and sublimations. There is no return of the repressed, as there is in neurotic symptoms. Defense mechanisms can, however, become character traits.

Although Freud is not explicit on this point, it may be surmised that he believed that character traits could represent either defenses against anxiety or rational solutions by the ego in dealing with its environment and with the id and superego. Again we are faced with the problem of how Freud would distinguish between the pathological and the normal in personality.

SOCIAL PSYCHOLOGY

Freud's contributions in the general area of social psychology may be summarized under five major headings: (1) socialization of the individual, (2) family structure and dynamics, (3) group psychology, (4) origin of society, and (5) the nature of human culture.

SOCIALIZATION OF THE INDIVIDUAL

Most of Freud's views on the socialization of the individual have already been discussed in the section on the superego, and not a great deal more can be added here. The child early in life becomes assimilated into the social tradition of his particular culture by the process of identification with his parents, and this assimilation is strengthened later in life by identification with other representatives of the culture (Freud, 1921). The agents of social tradition use rewards and punishments to coerce the child into acceptance of the cultural rules and customs (Freud, 1930). An intrapsychic system, the superego, takes over the role of external authority figures and exacts conformity to society. The person is then said to be socialized.

The person may rebel as well as conform to a social tradition. Rebellion may be produced by a number of factors, but two in particular are of major importance. These are (1) the severity of the hardships imposed by society on the individual and (2) the weakness of the ego and superego. If the prohibitions of society relative to its rewards are excessive, or if the person does not have the ability to find satisfactions within the framework of society (ego weakness), or if the social tradition has not been sufficiently incorporated due to a failure of identification (superego weakness), then the tendency to rebel against society is increased (Freud, 1927a). Crime is the most common expression of this revolt.

FAMILY STRUCTURE AND DYNAMICS

Freud has been criticized repeatedly for overemphasizing the role of biological determinants and underemphasizing the role of sociocultural influences in shaping the course of personality development. It is true that Freud does not give much consideration to the ways in which specific economic, political, religious, and social institutions may affect personality or how different types of social organization may produce different types of character, although in his book *Civilization and Its Discontents* (1930) he discusses at length how society imposes prohibitions upon the person and how the person may react to these prohibitions. One may, of course, apply these general principles in making an analysis of the influence of a particular institution or a particular form of society upon the people living in that society, as Kardiner (1939, 1945) has done. Freud himself has applied psychoanalytic concepts in the analysis of one important social institution, namely, religion. In *The Future of an Illusion* (1927a) and later in *Moses and Monotheism* (1939) he addresses himself to three major questions: (1) how religion originated, (2) how it has evolved, and (3) what effects it has upon the person. Yet the general impression which these books leave upon the reader is that in them Freud is writing from the standpoint of social criticism and not from the standpoint of social science. Freud, of course, was not a social scientist, and when he turned his attention to social phenomena it seemed to be with the intention of indicting society for its deleterious influences upon personality. The view of society that one gets from neurotic patients is not likely to be a favorable one, and this fact probably made it difficult for Freud to maintain a dispassionate attitude toward society. Moreover "society" did not treat Freud well in the early years of psychoanalysis, nor later when the Nazis came to power. Freud's general pessimism regarding society and its institutions has been remarked on more than once.

There is yet another social institution to which Freud has devoted an enormous amount of attention. This is the family. Freud's analysis of the structure and dynamics of the family as a primary group constitutes one of his greatest scientific achievements, and his most notable contribution to social psychology.

Let us first consider a three-member family, mother, father, and child. A relationship between the mother and father exists before the child is born, and the nature of this relationship will have a bearing upon the child's place in the family structure. The child's first libidinal object choice is the mother, while the father is seen as an interloper and rival. This triangular relationship involves the child in his first love and hate experiences. Since by the very nature of the case, he cannot maintain this pure dichotomy of feeling for the parents, modifications in the social triad occur. The mother disappoints the child by being an imperfect lover. She scolds and rejects him. She makes him do things he does not want to do and forbids him to do some of those things he would like to do. Consequently, hostile feelings are engendered which, existing side by side with love, produce ambivalence of feeling. Ambivalence appears to be the normal state of affairs in all close social relationships, the prototype of which is the mother-child relation. Likewise, hostility for the father does not continue to persist in an unchanged form. The boy makes an identification with the parent of the same sex, while the girl begins to transfer some of her love for the mother to the father. The reason for this transfer, in addition to possible biological ones, is that the mother is a source of greater disappointment to a daughter than to a son. The daughter, discovering that the mother does not have a penis, places the blame upon the mother for her own lack of a penis. She turns to the father, who possesses that which she wants, and at the same time her identification with the mother is strengthened. Later the girl's wish for a penis will become transformed into a wish for a baby (Freud, 1940). Thus, psychological processes play into the hands of biological functions.

The boy's disappointments with the mother are not nearly so great, because he does have a penis; yet this very fact constitutes a danger. He may be deprived of it by his father as a punishment for loving his mother too ardently. Fear of castration brings the oedipus complex to an end in the boy, whereas the female counterpart of the castration complex, penis envy, creates the oedipus complex in a girl (Freud, 1933). The girl's relationships with members of either sex always remain more complicated than do those of the boy.

Of course, the dynamics in any particular family may vary considerably from the foregoing model. Object choices and identifications, love, hostility, and ambivalence of feelings, the course of the oedipus complex and the intensity of the castration complex depend upon the personalities of the mother, father, and child and the ways in which their personalities interact with one another. Freud also places great emphasis upon the bisexual constitution of the child, and in one of his writings he wonders whether ambivalence toward the parents should not be attributed entirely to bisexuality (1923a).

The dynamics of a multiple-child family are similar to those of a single-child family. There is, however, one additional factor, namely, competition among the siblings for the exclusive possession of the loved parent and cooperation among them in expressing hostility against the resented parent. Since love and resentment are not divided between the two parents respectively, but instead there is always ambivalence of feeling toward both parents, the situation in regard to sibling competition and sibling cooperation is a complicated one.

The dynamics of family organization is reflected in the dynamics of all organized groups. For Freud, the family is the prototype for every association among people, whether the association be political, economic, social, recreational, or religious in character. The family is the bridge between individual psychology and social psychology.

GROUP PSYCHOLOGY

The essence of a group and that which holds it together are libidinal ties among the members. These ties are derived from aim-inhibited sexual impulses and from identifications which develop when object choices have to be renounced. Identification, therefore, is the key concept in Freud's theory of group dynamics. A primary group, according to Freud, is a number of individuals who have taken the same person, i.e., the leader, for their ideal and who, by virtue of having a common ideal, identify themselves with one another (Freud, 1921). Thus, emotional bonds are established among the members of a group and between each member and the leader. The bond with the leader is of more importance to group stability than are the ties between members, since it is the former which confers significance upon the latter. When the ties of the members with the leader are broken, the group dissolves unless another person acquires the status of an ideal figure.

The effects of the group upon the individual are to intensify his emotions and inhibit his intellectual functioning. These effects result from the libidinal relationships within the group, so that it may be said that in a group a person is governed more by his id than by his ego. He yields some of his ego autonomy for the sake of instinctual gratification (Freud, 1921).

The archaic prototype of a primary group is a father and his sons. The father forces sexual abstinence upon his sons, and the libido, denied discharge, is used to produce identifications between the sons and the father and among the sons. Freud says, "His (the father's) sexual jealousy and intolerance became in the last resort the causes of group psychology" (1921, p. 94). Group formation is a revival of a primal horde, in which the primal father is the group ideal which governs the ego in place of the ego ideal. Freud observes that the direct expression of sexual impulses is unfavorable to the formation of groups, although he observes that it is a remarkable fact, for which he does not know the explanation, that homosexuality even when it is overtly expressed is more compatible with the maintenance of group ties than is heterosexuality (1921).

Each individual is a member of numerous groups, so that his ties of identification take many directions and his ego ideal is a reflection of various leaders (Freud, 1921). (For a much fuller account of Freud's group psychology the reader is referred to Scheidlinger, 1952.)

ORIGIN OF SOCIETY

Freud's theory of the origin of society is set forth in detail in *Totem and Taboo* (1913) and is summarized in *Moses and Monotheism* (1939). In the dawn of human society, man lived in small hordes under the domination of a strong male. This strong male was the father of the whole horde. He had unlimited power which he used sadistically against his sons. Since all of the females of the horde were his property, the father could and did force the sons into sexual abstinence. If they excited his jealousy, they were killed or castrated or driven out of the tribe. Consequently, the sons were forced to steal their wives from other tribes, and this led to intertribal warfare, which further strengthened group ties within the horde. One of the sons, often the favored younger son, was chosen by the father to be his successor.

The sons who were driven out of the tribe organized and overcame the father. After killing the father, the sons ate him in order to produce identification with the father by incorporating a part of him.

Following this patricide, the sons began to quarrel with one another for power. Upon discovering that fraternal aggression was dangerous they formed the first social contract which was based upon a renunciation of instinctual gratification. To ensure group solidarity they established the taboo against incest and the law of exogamy. Thus the first form of society developed out of the family.

Brotherly love was instituted to control aggression among members of the same society, and aggression was drained off to some extent by declaring war on alien societies. War serves the purpose of keeping the in-group solidified.

NATURE OF CULTURE

Freud defines *culture* or *civilization*, terms used interchangeably in translating the German word *Kultur*, as "the sum of achievements and institutions which differentiate our lives from those of our animal forebears" (1930, p. 49). Culture serves two functions, it protects people against the dangers of the natural world, and it regulates the relations of human beings among themselves so that they do not destroy one another. Human existence in communities only becomes possible when a number of men unite together so that their strength is superior to that of any single individual. The substitution of the power of a united group of individuals for the exclusive domination of a single man is, in Freud's opinion, the decisive step toward civilization (Freud, 1930).

However, since a group can become united only by renouncing the freedom of individual instinctual gratification, civilization is necessarily frustrating. One gives up pleasure for security. However, the masses have no love for instinctual renunciation, so they must be coerced into working by an educated and intelligent minority. Freud wonders whether a change in society and universal education can convert the hostility of the masses toward culture into acceptance of it. Freud doubts whether they can, although he considers it a "noble project." Freud's pessimism grows out of his belief that people in general are not capable of extensive sublimation of aggressive and sexual impulses (Freud, 1930).

Freud is scornful of the popular notion that freedom is one of the major benefits of civilization, for, as he observes, civilized society deprives man of freedom instead of conferring it upon him. Society consists largely of imposing prohibitions upon a man's freedom of action. The ideal of society is not liberty but equality of treatment for all members of the society — in a word, justice. An unjust society, one in which the privileged few are accorded special favors while the many are held to a different standard of conduct, is bound to be an unstable society, since the masses will tend to rise up against the tyranny of injustice (Freud, 1930).

There is a similarity between cultural development and libidinal development. In both, the instincts are modified by the emergence of the reality principle as the dominant force in the individual and the group, by the formation of a superego, which is represented in society by laws and ideals, and by the institution of various defenses on a group as well as on an individual basis. Aggression and not sexuality is the most powerful obstacle which stands in the way of cultural progress. The aim-inhibited traits of tenderness and friendliness work in the interest of cementing cultural ties among people, while the aggressive impulse can give rise to no such positive contributions to cultural progress. Aggression can be drained off in intercultural hostility, but this is hardly a satisfactory solution, since the aim of warfare is always to destroy culture. Freud views the evolution of culture as a struggle between Eros and death (Freud, 1930).

It is paradoxical that culture, which was instituted to reduce human misery, should itself produce so much suffering. People are likely to overlook the benefits which culture bestows and to build up a tremendous amount of resentment and hostility against the self-abnegations imposed by culture. This hostility against culture expresses itself in envying primitive people who appear to be happier than civilized man (but who really are not, says Freud — they are probably worse off) and in a desire to regress to their level (Freud, 1930). Resentment against civilization also manifests itself in the otherworldliness of religion, as well as in many other forms, notably crime and antisocial actions. Freud also recognizes the role that civilization plays in exacerbating mental disorders (Freud, 1908b).

Recent Trends in Psychoanalytic Theorizing

EGO THEORY

The prevailing interest among psychoanalytic theorists today is the elaboration and reformulation of Freud's theory of the ego, which he enunciated in *Group Psychology and the Analysis of the Ego* (1921), *The Ego and the Id* (1923a), and later writings. As Anna Freud observed, in reference to these publications of her father, "the odium of analytical unorthodoxy no longer attached to the study of the ego and interest was definitely focused on the ego-institutions" (1936, p. 4). Theorists have made the most of this explicit sanction to study the ego, and the literature on this topic is an expanding one. The direction is away from studies of the defensive functions of the ego to studies of the synthetic (Nunberg, 1931), integrative (French, 1952), organizing (Hartmann, 1950), mastering (Hendrick, 1943), and multiple (Waelder, 1936) functions of the ego. Rapaport (1950, 1951a, 1951b, 1951c) has made a number of important contributions to ego theory, especially in the domain of thought processes. Hartmann, Kris, and Loewenstein (1947) have insisted that the ego should be defined in terms of its functions, and their emphasis has been more upon the healthy, adjustive, rational processes of the ego than upon its pathological formations. For these writers, however, ego defenses are not necessarily pathological or negative in character but may serve a healthy purpose in the formation of personality. "What developed as an outcome of defense against an instinctual drive may grow into a more or less independent and more or less structured function. It may come to serve different functions, like adjustment, organization, and so on" (Hartmann, 1950). These theorists have also revised Freud's basic hypothesis that the ego develops out of the id. They believe that both the ego and the id develop out of an undifferentiated phase which consists of a "matrix of animal instinct" and that consequently the ego is a partially independent unit of the personality, using neutralized libidinal and aggressive energy (Hartmann, 1952).

This trend of treating the ego as an autonomous system whose origin is parallel with the id and which is endowed with autonomous functions and energy sources has not gone unchallenged by other psychoanalysts. Nacht (1952), for example, deplores this new "ego psychology," which he considers sterile and regressive. He points out the dangers of "byzantinism" which threaten current psychoanalytic thinking and asserts that "Psychoanalysis has nothing to gain from these exegetic elaborations in which the letter always takes precedence of the spirit" (p. 50).

The development of psychoanalytic ego psychology is seen by one writer (French, 1952) as offering an opportunity to effect a *rapprochement* between psychoanalysis and psychology. French himself has attempted to integrate the views of Tolman, Hull, Pavlov, and Lewin with those of psychoanalytic theorists.

For social psychology Erikson's formulations regarding the ego are probably the most relevant. He characterizes his book, *Childhood and Society* (Erikson, 1950), as being in a sense "a psychoanalytic book on the relations of the ego to society" (1950, p. 12). In this book he tries to show by case studies of people living in different circumstances and in different cultures how the ego's development is inevitably bound up with the nature of social organization. He has also demonstrated how the ego's development is related to changing historical reality (Erikson, 1947). Erikson's principal concepts are those of ego identity and group identity. In its subjective meaning ego identity is an awareness of the fact that there is continuity in the ego's synthesizing methods and the feeling that these methods are effective in safeguarding the sameness and continuity of one's meaning for others. Group identity is the group's basic ways of organizing experience for the members of the group. In general, Erikson has had more to say about society's effect upon the person than about the person's effect upon society.

INSTINCT THEORY

Virtually all theoretical psychoanalysts reject or ignore Freud's death instinct. Hartmann, Kris, and Loewenstein (1949), for example, assume an innate aggressive or destructive drive which is independent of a death instinct. For these writers, sex and aggression are the two basic instincts or "instinctual drives," as they prefer to call them, since the term "drive" connotes greater plasticity than the term "instinct" does. (See also Hartmann, 1948.)

Simmel (1944), on the other hand, believes that aggressive and destructive behavior is a manifestation of self-preservation or, more specifically, of eating. Libido is the common energy for both sex and aggression.

More and more, instinct theory is losing popularity among contemporary theorists and is being handed over to the biologists, since it is felt that instinct is a biological concept. This trend is but a part of the larger movement to "debiologize" psychoanalysis by minimizing the role of heredity and maximizing the role of society in the formation of personality. Karen Horney (1939) represents an extreme proponent of this movement.

Another feature of contemporary psychoanalytic theorizing is the substitution of ontogenetic propositions for phylogenetic propositions. Hartmann and Kris (1945) attribute Freud's predilection for "acquired characters" to the biological *Zeitgeist* of the nineteenth century in which he was reared. They do not believe that there is sufficient empirical support for the Lamarkian explanation of evolution. For Hartmann and Kris, fear of castration is "in the air" rather than in the genes. Curiously enough, it seems not to have occurred to psychoanalysts that it is possible to substitute a phylogenesis based upon mutation and natural selection for one based upon acquired characters.

RESEARCH METHODS

For validating their genetic hypotheses, psychoanalytic investigators have consistently favored the method of reconstruction. The past of the person, who is usually in the role of a patient, is reconstructed from his free associations and dreams, yielding causal connections between childhood events and adult formations. Within recent years another method, that of discovering the origins of personality by observing infants and children, has become popular. The results of many of these investigations have been appearing in the annual volumes of *The Psychoanalytic Study of the Child.*

Psychoanalysts have also been urged to make more use of experimental methods (Kubie, 1952) and comparative studies with animals (Levy, 1952), but as yet these promptings have not borne much fruit. The experimental validation of psychoanalytic propositions has been pursued primarily by nonpsychoanalysts (Hilgard, 1952). Research of special relevance to psychoanalytic theory has been surveyed by Sears (1943, 1944) and more recently by Hilgard (1952), Kubie (1952), and Blum (1953).

Application of Psychoanalytic Theory in the Social Sciences

In the judgment of many thoughtful people, Freudian thought is one of the main currents in the intellectual stream of modern life. There is scarcely any aspect of the contemporary scene that has not felt the impact of Freud's thought. The social sciences are no exception.

It would be a tremendous and unrewarding task to search out and document every application of Freudian concepts in the social sciences, since so many of the applications have been made without noting or perhaps without knowing their source. These hidden influences we shall not take into account. Instead we shall limit our analysis to those self-conscious attempts to apply Freudian theory in the social sciences. It is well to bear in mind that some of the contributions have been made by professional social scientists, while others have been made by professional psychoanalysts who have become interested in a particular social science. There are drawbacks on both sides. The social scientist is not likely to have intimate and personal knowledge of the clinical context from which psychoanalytic theory has sprung, and he is less likely to view psychoanalysis in the full scope of its development than is the practitioner of this discipline. The psychoanalyst works under the handicap of being an amateur in the field to which he is trying to apply his concepts, and amateurs are prone to make errors and also are likely to be viewed with suspicion by professionals.

ANTHROPOLOGY

In relative amount the impact of psychoanalysis upon cultural anthropology has probably been greater than upon any other social science discipline. Kluckhohn, in discussing the relation between psychiatry and anthropology, suggests that: ". . . American anthropology, for good or for ill, has seemed to find only in psychoanalysis the bases for a workable social psychology" (1944b, p. 590). He notes further that, with the exception of relatively recent contact with Horney, Kardiner, and Fromm, almost all of the influence has stemmed directly from orthodox Freudian psychology.

The first evidence of anthropological interest in psychoanalysis did not materially precede similar signs in sociology and psychology, but the relative importance of this interaction eventually became very different. Kluckhohn (1944b) suggests that interest in psychoanalytic theory on the part of an American anthropologist (ten years after similar recognition on the part of the English anthropologists, Rivers, 1926, and Seligman, 1932) was first evident in a note published by E. Parsons (1915) in a psychoanalytic journal. It was not until 1920 in a paper by Boas that any important notice was taken of psychoanalysis, and even here Boas' suggestion that there was merit for the anthropologist was surrounded by much more caution than enthusiasm. At about the same time, Kroeber (1920) reviewed Freud's *Totem and Taboo* (1913) in a highly critical vein, but nevertheless indicated there was much of value in the volume for the professional anthropologist. Some twenty years later Kroeber suggested that he still felt his criticisms were warranted but went on to say, in connection with the original review:

> . . . I found myself somewhat conscience-stricken when, perhaps a decade later, I listened to a student in Sapir's seminar in Chicago making his report on *Totem and Taboo,* who, like myself, first spread out its gossamer texture and then laboriously tore it to shreds. It is a procedure too suggestive of breaking a butterfly on the wheel. An irridescent fantasy deserves a more delicate touch even in the act of demonstration of its unreality. (1939, p. 446)

In spite of these initial signs of awakening interest in Freud there seems little question but that the early attitude of anthropologists toward psychoanalysis was one of firm hostility. This initial view was sufficiently marked so that both Sapir (1934) and Kluckhohn (1944b), following him, suggested the possible influence of personal factors in the professional anthropologist, which had led him into his profession in the effort to escape from the personal and familiar world and which also served to produce considerable resistance to the highly personal and psychological theories of psychoanalysis. The determinants that Sapir points to are, of course, double-edged, the same factors that lead a person to seek depersonalized experience may cause him to become violently interested in intimate and highly personal experience. This reversibility may account in part for the extent to which psychoanalysis later became a focus of interest for many social anthropologists.

Among the first of the American anthropologists to show the influence of Freud were Sapir, Malinowski, and Mead. All three of these figures were highly skeptical of the general merit of psychoanalysis as a tool for the ethnologist. In spite of their unwillingness to embrace psychoanalytic theory, however, they were heavily influenced by psychoanalytic assumptions and hypotheses in their field work and publications. Prior to the thirties there was little attempt on the part of the ethnologist to publish the details of individual life histories or the developmental process — material of crucial significance to the social psychologist and psychoanalyst (Kluckhohn, 1944b, 1945). There seems little doubt that the changes in anthropological modes of field work and publication practices were closely related to an increasing interest in psychological problems, in part stimulated by Freud and his theory. The increased interest in this kind of data that was manifested by both psychoanalysts and anthropologists during the thirties led directly to the formation of a small but vocal group of "personality and culture" students. Individuals interested in this borderline area remain a potent force within contemporary anthropology, and the effectiveness of their efforts is demonstrated by many of the studies discussed in Chapters 25 and 26.

We might roughly characterize the influence of psychoanalysis upon anthropology in the following way. First, we find anthropologists working vigorously to demonstrate the fallibility of formulations involving universal psychological processes. Thus, Malinowski (1927) consistently and Mead (1928, 1932, 1935), at least early in her career, appeared motivated to provide data in order to explode the myth of psychoanalysis. Second, there were anthropologists who showed an interest in psychoanalysis as a means of accounting for puzzling aspects of the data they observed and who also tended to shift their field-work tactics somewhat to produce data that were more pertinent to psychoanalytic propositions (Kluckhohn, 1942, 1944a, 1947, 1951; Opler, 1935, 1936, 1938; Beaglehole, 1937, 1938). Third, we have a number of examples of the fruitful collaboration of anthropologists and psychoanalysts in investigation, or at least in the analysis of data. Easily the outstanding illustration of this trend is the work of Linton and Kardiner (Linton, 1945; Kardiner, 1939, 1945), although there are numerous less well-known examples, many of them discussed by Kluckhohn (1944b). The Columbia University Research in Contemporary Culture, described by Mead (1951), is perhaps the most ambitious present-day attempt at the interdisciplinary application of psychoanalytic theory to cross-cultural data. Fourth, and overlapping with the previous category, are the anthropologists who have made serious attempts to equip themselves to collect and interpret data that are especially relevant to psychoanalytic theory (DuBois, 1944; Henry, 1940, 1941; Henry and Henry, 1953; Hallowell, 1936, 1941, 1942, 1951). Fifth, there are the small group of psychoanalysts (Roheim, 1934, 1943, 1947, 1950a, 1950b, 1951; Erikson, 1939, 1950; see also Wilbur and Muensterberger (Eds.), 1951) who have actually conducted field-work investigations of their own and interpreted their findings within a psychoanalytic framework.

SOCIOLOGY

The impact of psychoanalysis upon sociology has been considerably less than upon anthropology. Devereux (1952) suggests that this state of affairs has arisen in spite of the greater relevance of psychoanalysis for the problems of the sociologist than for those of the cultural anthropologist. This is a somewhat moot point, as the conceptual differences between sociology and anthropology are difficult to specify and a matter for argument even among experts. However, it seems safe to conclude that psychoanalysis is *as* relevant for the problems of the sociologist as for those of the anthropologist. If this is the case, why is it that the two fields have been influenced to such different degrees? In a general discussion of the influence of psychoanalysis upon sociology, Burgess (1939) has considered some of the factors that may have dampened the enthusiasm of sociologists for psychoanalysis. Among other factors, he suggests the following as responsible for the sociologist's rejection of psychoanalysis: objection to representing all behavior as motivated by sexual needs, a preference for accounting for behavior in terms of cultural determinants, the failure of psychoanalysis to integrate with previous instinct theories and in addition a general trend away from instinct concepts on the part of sociologists, the appearance of deviant psychoanalytic groups, and the preoccupation of sociologists with their own theories and problems. While some of these factors undoubtedly played a role in dulling the interest of the sociologist in psychoanalysis, few of them seem equal to accounting for the difference in the reaction of the sociologist and the anthropologist, for most are equally applicable to both groups. Only in the case of the move away from instinct theories does there appear to be an important difference between the two disciplines. And even here this is apparent primarily because the leaders in the anti-instinct movement were sociologists (Bernard, 1921; Faris, 1921).

If Burgess' formulation does not differentiate between the anthropologist and the sociologist, what does? First of all it is clear that historical accident must have played some role here, as in most complex phenomena. The fact that Sapir, Malinowski, and Mead were influenced by psychoanalysis is not alone tied to the fact that they were anthropologists. There must also have been some accidental factors linked to the particular experiences they were exposed to and the kinds of people they were. More important, the anthropologist, by virtue of his field-work experiences, was in much more intimate contact with the details of human existence than the sociologist. This means that he was exposed to data relevant to psychoanalytic formulation and also that he was more inclined toward acceptance of the methods of the psychoanalyst, as they resembled his own. Third, by the second decade of this century there was a vigorous positivistic movement in sociology, with its accompanying stress upon quantitative data and careful empirical controls. Certainly, for the demographer the subjective and poorly specified data of the analyst must have appeared a long step backward in the attempt to develop a sound discipline concerned with social behavior. Among the anthropologists there was no equivalent affection for quantification and "rigor." Fourth, American anthropology under the influence of Franz Boas was much more theory-less than American sociology and there was consequently less likelihood that psychoanalysis would be rejected because of conflicting or competing theoretical convictions. Fifth, sociology in the early decades of this century is generally considered to have been tinged with the goals and values of the Protestant ministry. From such a vantage point Freud's pessimistic and sensate conception of man was certain to be greeted with hostility. Sixth, psychoanalytic formulations were directly concerned with many classical problems of anthropology (ritual, incest taboo, kinship patterns, marriage practices), while these same topics were of only secondary interest to the sociologist. And finally, it is evident that Freud's early excursions into areas specifically labeled anthropological, and subsequent applications by Jones (1923, 1924, 1925) and Reik (1928), made it more difficult for the anthropologist to ignore psychoanalysis than the sociologist.

Granted its slowness to respond to psychoanalytic theory, what is the current state of sociology in this regard? What are the traces of psychoanalytic influence? The impact has been most continuously evident in the area of the family where, understandably, psychoanalytic formulations have had a considerable vogue. From the mid-twenties, writers in this area have dealt more and more heavily with conceptions tinged with Freud's view of the family and the process of socialization. The operation of this influence during the thirties is well illustrated by the volumes of Folsom (1934) and Waller (1938). More recently, the influence is clearly shown in the text written by Burgess and Locke (1945) and culminated in the volume by Winch (1952). In the latter we find both formulations and actual investigations carried out by the author heavily indebted to psychoanalysis. Perhaps more impressive is the extent to which psychoanalytic formulations appear to have influenced sociological theory. Easily the best example in this area is the recent work of Talcott Parsons (1947, 1950, 1951, 1952a, 1952b, 1953) who, in his persistent attempts to develop a general theory of social action, has leaned heavily upon psychoanalytic concepts in the effort to represent the motivational process. The collaboration of Parsons, Shils, Tolman *et al* in the attempt to develop an interdisciplinary theory of behavior (Parsons and Shils, 1951) drew heavily from psychoanalytic sources. Interesting papers by Green (1945, 1946), Davis (1936), Shils and Janowitz (1948), Devereux (1939a, 1939b, 1942), and Bales (1945, 1946), as well as papers in the volume honoring Roheim (Wilbur and Muensterberger, 1951), provide further evidence of growing interest on the part of sociologists in conceptions influenced by psychoanalysis. A final important example of the influence of Freud upon sociological thought is found in the works of David Reisman (1950a, 1950b, 1950c).

SOCIAL PSYCHOLOGY

In 1932 Fay Karpf (1932) surveyed the history and current status of social psychology in the United States, and the European background from which it developed. Her book does not indicate in any way whatsoever that Freud has had the slightest effect upon the development of social-psychological thinking. Nearly twenty years later, Krech, in his survey of psychological theory and social psychology (1951), writes

> . . . it is obvious that much of the current vitality of social psychological theory is due to the influence of psychoanalysis and of those social psychologists trained in, and familiar with, the concepts of psychoanalysis. (p. 680)

Certainly Krech's point of view is fashionable among contemporary social psychologists. Evidence of this is provided by the Lewin Memorial Address of 1948 in which both speakers (Allport and Tolman) alluded to Freud and Lewin as the most important sources of contemporary influence in this area of psychology. If it is true that some time during the past two decades psychoanalysis made its impact upon social psychology, we should expect to find evidence of it in textbooks and reported investigations within social psychology.

A sampling of textbooks written during the thirties reveals that psychoanalytic theory, with one exception, finds little application or systematic attention relative to the problems of social psychology. Freud is usually mentioned, and sometimes aspects of his general theory, e.g., wit, sublimation, taboo, conflict, are presented, but often his theories are misrepresented, and always the integrity of psychoanalytic theory is damaged by referring to concepts out of context. The one exception to the rule is J. F. Brown's *Psychology and the Social Order* (1936). It is obvious that Brown is well grounded in psychoanalytic theory as well as in field theory, and his serious effort to synthesize these two major viewpoints is without parallel in social psychology. He is one of the few social psychologists who has paid much attention to Freud's social psychology as set forth in *Group Psychology and the Analysis of the Ego* (1921), while his chapter on the structure and genesis of personality is, for its day, a succinct and fairly accurate resumé of psychoanalytic theory. Like most social psychologists, Brown believes that Freud overemphasized the role of biological determinants in personality formation and social interactions, and he is critical of Freud's methodology. These latter criticisms are presented in more detail in Brown's article on *Freud and the Scientific Method* (1934). During the thirties, then, it may be said that only one out of many social psychology textbook writers reacted to Freud in any significant way. Nor does a perusal of the first handbook of social psychology (Murchison, 1935) change the picture. With the exception of two chapters, Willoughby's on magic and Wells' on adaptive regression, Freud and Freudian concepts are rarely mentioned.

The social psychology textbooks of the forties do not change the picture. Freud is still mentioned, and some of his concepts and special theories are presented briefly, but there is little evidence that Freud has made much of a contribution to current social psychological thought. In view of Krech's opinion quoted above, one might expect to find much psychoanalytic theory incorporated in Krech and Crutchfield's influential book, *Theories and Problems of Social Psychology* (1948), but such is not the case. There is some talk about unconscious motivation and other psychoanalytic concepts, but there is no mention of Freud's group psychology. The vitality of this book, as so many other recent volumes in social psychology, stems from the theories of gestalt psychology, Lewin, and Tolman, not from psychoanalysis.

Three recent textbooks have been surveyed. Queener's *Introduction to Social Psychology* (1951) observes that there is a great deal of irrational behavior in social life, the understanding of which is provided for only by psychoanalytic theory. Yet there is little evidence in this book that psychoanalytic concepts have been used for understanding anything. Asch's *Social Psychology* (1952) is written entirely from a gestalt viewpoint and is critical of Freud's negative and pessimistic attitudes toward man and society. Asch asserts that Freud ignored the productive forces of man, that he overemphasizes social friction, and that he gives a one-sided picture of the oppressiveness of society on the individual. Nevertheless, Asch is of the opinion that the guiding ideas of social psychology are derived from behaviorism and psychoanalysis. Newcomb's *Social Psychology* (1950) contains more references to Freud than to any other single person, and his presentation of psychoanalytic theory is more complete than that of any other textbook in social psychology since J. F. Brown's, although oddly enough he does not refer to Freud's *Group Psychology and the Analysis of the Ego*. Yet when one examines the substance of what Newcomb has to say about psychoanalysts two features stand out. The revisionists, Fromm, Horney, and Sullivan, contribute much more to Newcomb's theory of social interaction than does Freud, a fact that Newcomb frankly acknowledges. The second feature is that Freud has had much less influence upon Newcomb than Newcomb has had upon Freud. What Newcomb does repeatedly is to translate psychoanalytic theory into Newcomb's version of role theory. As Newcomb says, "Many of Freud's contributions — like those of other really great men — stand up under the necessity of being translated into the language of a later day" (p. 376).

This survey of textbooks of social psychology fails to support Krech's opinion regarding the vital influence of psychoanalysis upon social psychology. Revisionists like Fromm and Horney, who have placed the main accent on cultural determinants of behavior, have had considerably more influence than Freud probably because their views are more in accord with the dominant stream of social psychological thought in this country. Perhaps one should not say that they have had an influence, since they have chosen to swim with the current rather than to make new ones. They have been more influenced than influential.

When we turn from a survey of textbooks to a survey of current research trends in social psychology the situation is somewhat different. Much of the research in social psychology is, of course, largely empirical and atheoretical in character. That which is theory oriented derives most of its testable hypotheses from gestalt psychology, Lewin, role theory, perception-cognition theory, and neobehaviorism. This impression is fortified by examination of the recent edition of *Readings in Social Psychology* (Swanson, Newcomb, and Hartley, 1952). This volume intends to present a representative picture of the research in social psychology, and a large and diverse committee aided in the selection of the studies to be included. Freud's name is mentioned only four times in the index and in two of these cases the reference is to a critical statement. More significant is the incidence of studies that can be seen as growing out of psychoanalytic theory. Even using very generous standards not more than a half dozen papers seem in any direct way related to psychoanalytic theory. This is in striking contrast to Lewin's field theory.

Psychoanalytic theory, however, has made some significant contributions to social psychological research. One example of this is found in the area of prejudice. The authors of the widely known volume, *The Authoritarian Personality* (Adorno, *et al.* 1950) acknowledge their debt to Freudian theory in formulating their research hypotheses. Other writers (Ackerman and Jahoda, 1950; Loewenstein, 1951; Lowenthal and Guterman, 1949; Simmel (Ed.) 1946) have used the insights of psychoanalysis to illuminate the roots of prejudice. The well-known scapegoat theory of prejudice represents a relatively straightforward application of psychoanalytic theory (Freud, 1939) and has led to several investigations (Miller and Bugelski, 1948; Lindzey, 1950). Psychoanalytic theory has also been applied in the area of delinquency (Aichorn, 1935; Eissler (Ed.) 1949; Friedlander, 1947). Research in social psychology being conducted at the Tavistock Institute of Human Relations in London is heavily influenced by psychoanalytic theory. Especially noteworthy are the studies of W. R. Bion on experiences in a group (Bion, 1948a, 1948b, 1949a, 1949b, 1950a, 1950b, 1951). Waelder's (1939) discussion of the conditions leading to war is a notable application of orthodox psychoanalytic theory to a complex social phenomenon. Numerous studies concerned with the socialization process and child development reflect psychoanalytic thinking, as the literature summarized in Chapter 18 clearly indicates. Bowlby's (1951) survey of studies of the effect of maternal care upon mental health provides evidence of the vitality of psychoanalytically influenced studies in this area. The recent increase in studies concerned with social-psychological problems in a cross-cultural frame of reference also shows some attention to Freudian formulations (McClelland and Friedman, 1952; Chapters 14 and 18, this work). The avalanche of investigation that followed publication of *Frustration and Aggression* (Dollard, Doob, *et al,* 1939) can be considered heavily indebted to Freud's ideas, as the authors of this monograph indicate. The impact of psychoanalysis has been felt to some extent in the study of propaganda or mass media, as studies by Kris and Speier (1944) and Kecskemeti and Leites (1947, 1948a, 1948b, 1948c) illustrate. In the investigation of small group behavior the work of Bales (1950) and Miller and Stine (1951) represent attempts to derive insights from psychoanalysis. Finally, psychoanalysis has had a great deal of indirect influence upon research in social psychology through the medium of Lewin and his students. A number of Freudian concepts, prominent among them displacement and conflict, have been adapted and incorporated in Lewin's own conceptions and in this form have had heavy impact upon contemporary research and conceptions of social psychologists.

Are there events that have served to lessen the interplay between psychoanalysis and social psychology? Certainly one important factor is that psychoanalysis is regarded as an "instinct" psychology, and social psychologists have been opposed to instincts since the twenties. Another clearly significant factor is that Freud's pessimism regarding society is incompatible with the environmentalism and meliorism among social psychologists, and their conviction that man can find happiness in society as soon as certain attainable changes are made in society. A third factor is that psychoanalysis has been closely associated with psychopathology and for the most part psychoanalytic theory has not been presented in such a way that the social psychologist could easily see the relevance of it for his field. Probably a fourth consideration is that psychoanalytic theory has not been adequately taught in American universities. One cannot apply something of which one is ignorant. The inadequacy of instruction is associated with the open hostility of many psychologists toward psychoanalysis, as well as the strong in-group tendencies on the part of most psychoanalytic groups. Marjorie Brierley (1951), an English psychoanalyst, believes that the potential contribution of psychoanalysis to social psychology is very great but it is not being realized because "disabling tendencies to fission and isolation" (p. 153) exist in psychoanalysis.

OTHER SOCIAL SCIENCES

Aside from occasional instances, psychoanalytic theory has not penetrated extensively into the domains of history, economics, political science, and law, although it has had considerable influence upon the humanities. The early and important writing of Harold Lasswell (1927, 1930, 1935, 1948) and the book of readings edited by Stanton and Perry (1951) show the influence of psychoanalysis on political science. In law, the volume by Jerome Frank (1930) depends almost exclusively on Freudian theory. In economics there is practically nothing except Schneider's (1948) effort to compare Freudian psychology with Veblen's social theory, while in history the influence of Freud is difficult to evaluate, although biographers have been using psychoanalytic concepts in recent years to explain the conduct of their subjects. De Saussure, a psychoanalyst, is amazed "that psychoanalysis has not been systematically applied by historians, because in itself the method is historical" (1950, p. 7). Psychoanalysts have shown no reluctance in applying their concepts and methods in the analysis of historical figures and movements. Freud did just this in *Leonardo da Vinci* (1910a) and in *Moses and Monotheism* (1939). Loewenfeld has remarked that "the role of the great man in history is a special field for analytic investigation" (1950, p. 289), and Fromm (1930, 1941) has applied psychoanalytic concepts freely in historical analysis.

SUMMARY

We have attempted to present a detailed summary of Freud's theory of personality and its application to social phenomena. In addition, we have tried to summarize briefly some of the major features of the influence psychoanalysis has exerted upon the social sciences. Our emphasis upon the structure of Freud's theory was determined largely by the fact that this theory, especially as it pertains to the psychology of personality and to social psychology, has not always been clearly understood by psychologists. There are many reasons for this confusion, among them the sheer quantity of Freud's writings and the much greater popularity of Freud's theories of abnormal behavior. The complexity of this theory and limitations of space have made it impossible for us to deal with many issues of great importance in a full consideration of this topic. The contributions of Adler, Jung, and the neo-Freudians have had to be overlooked in spite of their favor among many social psychologists. Further, the many difficulties involved in the use of psychoanalytic theory as a guide to thought and research in this area, as well as the structural shortcomings of the theory itself, have not been mentioned. Our sole intention, then, has been to present to the reader the core of psychoanalytic theory and to convey some of our conviction that it is a potential tool of great power in the hands of interested psychologists.

REFERENCES

Ackerman, N. W., & Jahoda, Marie *Antisemitism and emotional disorder.* New York: Harper, 1950.

Adorno, T. W., Frenkel-Brunswik, Else, Levinson, D. J., & Sanford, R. N. *The authoritarian personality.* New York: Harper, 1950.

Aichhorn, A. *Wayward youth.* New York: Viking, 1935.

Asch, S. E. *Social psychology.* New York: Prentice-Hall, 1952.

Bales, R. F. Social therapy for a social disorder-compulsive drinking. *J. soc. Issues,* 1945, *1,* 14-22.

Bales, R. F. Cultural differences in rates of alcoholism. *Quart. J. Stud. Alcohol,* 1946, *6,* 480-499.

Bales, R. F. *Interaction process analysis.* Cambridge, Mass.: Addison-Wesley, 1950.

Beaglehole, E. Emotional release in a Polynesian community. *J. abnorm. soc. Psychol.,* 1937, *32,* 319-328.

Beaglehole, E. A note on cultural compensation. *J. abnorm. soc. Psychol.,* 1938, *33,* 121-123.

Bernard, L. L. The misuse of instinct in the social sciences. *Psychol. Rev.,* 1921, *28,* 96-119.

Bernfeld, S. Freud's earliest theories and the school of Helmholtz. *Psychoanalyt. Quart.,* 1944, *13,* 341-362.

Bion, W. R. Experiences in Groups: I. *Hum. Relat.,* 1948a, *1,* 314-320.

Bion, W. R. Experiences in Groups: II. *Hum. Relat.,* 1948b, *1,* 487-496.

Bion, W. R. Experiences in Groups: III. *Hum. Relat.,* 1949a, *2,* 13-22.

Bion, W. R. Experiences in Groups: IV. *Hum. Relat.,* 1949b, *2,* 295-303.

Bion, W. R. Experiences in Groups: V. *Hum. Relat.,* 1950a, *3,* 3-14.

Bion, W. R. Experiences in Groups: VI. *Hum. Relat.,* 1950b, *3,* 395-402.

Bion, W. R. Experiences in Groups: VII. *Hum. Relat.,* 1951, *4,* 221-227.

Blum, G. S. *Psychoanalytic theories of personality.* New York: McGraw-Hill, 1953.

Boring, E. G. *A history of experimental psychology.* 2nd. Ed. New York: Appleton-Century-Crofts, 1950.

Bowlby, J. *Maternal care and mental health.* Geneva: World Health Organization, *Monogr. 2,* 1951.

Brierley, Marjorie *Trends in psychoanalysis.* London: Hogarth Press, 1951.

Brown, J. F. Freud and the scientific method. *Phil. Sci.,* 1934, *1,* 323-337.

Brown, J. F. *Psychology and the social order.* New York: McGraw-Hill, 1936.

Burgess, E. W. The influence of Sigmund Freud upon sociology in the United States. *Amer. J. Sociol.,* 1939, *45,* 356-374.

Burgess, E. W., & Locke, H. J. *The family: From institution to companionship.* New York: American Book Co., 1945

Davis, K. Jealousy and sexual property. *Soc. Forces,* 1936, *14,* 395-405.

de Saussure, R. Psychoanalysis and history. In G. Roheim (Ed.), *Psychoanalysis and the Social Sciences.* Vol. 2. New York: International Universities Press, 1950. Pp. 7-66.

Devereux, G. A sociological theory of schizophrenia. *Psychoanal. Rev.,* 1939a, *26,* 315-342.

Devereux, G. Maladjustment and social neurosis. *Amer. sociol. Rev.,* 1939b, *4,* 844-851.

Devereux, G. Social structure and the economy of affective bonds. *Psychoanal. Rev.,* 1942, *29,* 303-314.

Devereux, G. Sociology. In J. Frosch (Ed.), *The annual survey of psychoanalysis.* New York: International Univ. Press, 1952. Pp. 325-362.

Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. R. *Frustration and aggression.* New Haven: Yale Univ. Press, 1939.

Du Bois, Cora A. *The people of Alor; a social-psychological study of an East Indian island.* Minneapolis, Minn.: Univ. of Minn. Press, 1944.

Eissler, K. R. (Ed.) *Searchlights on delinquency.* New York: International Univer. Press, 1949.

Erikson, E. H. Observations on Sioux education. *J. Psychol.,* 1939, *7,* 101-156.

Erikson, E. H. Ego development and historical change. In Anna Freud and others (Eds.), *The psychoanalytic study of the child.* Vol. 2. New York: International Univ. Press, 1947. Pp. 359-396.

Erikson, E. H. *Childhood and society.* New York: Norton, 1950.

Faris, E. Are instincts data or hypotheses? *Amer. J. Sociol.,* 1921, *27,* 184-196.

Fenichel, O. *The psychoanalytic theory of neurosis.* New York: Norton, 1945.

Folsom, J. K. *The family; Its sociology and social psychiatry.* New York: Wiley, 1934.

Frank, J. *Law and the modern mind.* New York: Brentano's, 1930.

French, T. M. *The integration of behavior.* Vol. I. Basic postulates. Chicago: Univer. of Chicago Press, 1952.

Freud, Anna *The ego and the mechanisms of defence.* New York: International Univer. Press, 1946. (First publishd in German in 1936.)

Freud, S. (1894) The defence neuro-psychoses. In *Collected Papers.* Vol. I. London: Hogarth Press, 1924. Pp. 59-75. (First published in the *Neurologisches Zentralblatt,* 1894, Nos. 10 and 11.)

Freud, S. (1900) The interpretation of dreams. In *The Basic Writings of Sigmund Freud.* New York: Random House, 1938. Pp. 181-549. (Translated from the 8th edition. *Die Traumdeutung,* 1932. First edition, 1900.)

Freud, S. (1904) Psychopathology of everyday life. In *The Basic Writings of Sigmund Freud.* New York: Random House, 1938. Pp. 35-178. (Translated from *Zur Psychopathologie des Alltagslebens,* Vienna, 1904.)

Freud, S. (1905) Three contributions to the theory of sex. In *The Basic Writings of Sigmund Freud.* New York: Random House, 1938. Pp. 553-629. (Translated from *Drei Abhandlungen zur Sexualtheorie,* Vienna, 1905.)

Freud, S. (1908a) Character and anal erotism. In *Collected Papers.* Vol. II. London: Hogarth Press, 1924. Pp. 45-50. (First published in *Psychiatrisch-Neurologische Wochenschrift,* Bd. IX, 1908.)

Freud, S. (1908b) "Civilized" sexual morality and modern nervousness. In *Collected Papers.* Vol. II. London: Hogarth Press, 1924. Pp. 76-99. (First published in *Sexual Probleme,* new issue of the periodical *Mutterschutz,* Bd. IV, 1908.)

Freud, S. (1910a) *Leonardo Da Vinci: A study in psychosexuality.* New York: Random House, 1947. (Translated from *Eine Kindheitserinnerung des Leonardi da Vinci,* Vienna, 1910.)

Freud, S. (1910b) Psychogenic visual disturbance according to psycho-analytical conceptions. In *Collected Papers.* Vol. II. London: Hogarth Press, 1924. Pp. 105-112. (First published in *Arztliche Standeszeitung,* Vienna, 1910.)

Freud, S. (1911) Formulations regarding the two principles in mental functioning. In *Collected Papers.* Vol. IV. London: Hogarth Press, 1925. Pp. 13-21. (First published in *Jahrbuch,* Bd. III, 1911.)

Freud, S. (1912) Types of neurotic nosogenesis. In *Collected Papers.* Vol. II. London: Hogarth Press, 1924. Pp. 113-121. (First published in *Zentralblatt,* Bd. II, 1912.)

Freud, S. (1913) Totem and taboo. In *The Basic Writings of Sigmund Freud.* New York: Random House, 1938. Pp. 807-930. (Translated from *Totem and Tabu,* Vienna, 1913.)

Freud, S. (1914) On narcissism: an introduction. In *Collected Papers. Vol. IV.* London: Hogarth Press, 1925. Pp. 30-59. (First published in *Jahrbuch,* Bd. VI, 1914.)

Freud, S. (1915a) Some character-types met with in psycho-analytic work. In *Collected Papers.* Vol. IV. London: Hogarth Press, 1925. Pp. 318-344. (First published in *Imago,* Bd. IV, 1915-16.)

Freud, S. (1915b) The unconscious. In *Collected Papers.* Vol. IV. London: Hogarth Press, 1925. Pp. 98-136. (First published in *Int. Z. Psychoanal.,* Bd. III, 1915.)

Freud, S. (1915c) Instincts and their vicissitudes. In *Collected Papers.* Vol. IV. London: Hogarth Press, 1925. Pp. 60-83. (First published in *Int. Z. Psychoanal.,* Bd. III, 1915.)

Freud, S. (1915d) Repression. In *Collected Papers.* Vol. IV. London: Hogarth Press, 1925. Pp. 84-97. (First published in *Int. Z. Psychoanal.,* Bd. III, 1915.)

Freud, S. (1916) Metapsychological supplement to the theory of dreams. In *Collected Papers,* Vol. IV. London: Hogarth Press, 1925. Pp. 137-151. (First published in *Int. Z. Psychoanal.,* Bd. IV, 1916-18.)

Freud, S. (1920) *Beyond the pleasure principle.* London: Hogarth Press, 1922. (Translated from *Jenseits des Lustprinzips,* Vienna, 1920.)

Freud, S. (1921) *Group psychology and the analysis of the ego.* London: Hogarth Press, 1922. (Translated from *Massenpsychologie und Ich-Analyse,* Vienna, 1921.)

Freud, S. (1923a) *The ego and the id.* London: Hogarth Press, 1927. (Translated from *Das Ich und das Es,* Vienna, 1923.)

Freud, S. (1923b) The libido theory. In *Collected Papers.* Vol. V. London: Hogarth Press, 1950. Pp. 131-135. (First published in *Handworterbuch fur Sexualwissenschaft,* 1923.)

Freud, S. (1924) The economic problem in masochism. In *Collected Papers,* Vol. II. London: Hogarth Press, 1924. Pp. 255-268. (First published in *Int. Z. Psychoanal.,* Bd. X, 1924.)

Freud, S. (1925a) Negation. In *Collected Papers.* Vol. V. London: Hogarth Press, 1950. Pp. 181-185. (First published in *Imago,* 1925. *Vol. II.*)

Freud, S. (1925b) A note upon the "mystic writing pad." In *Collected Papers.* Vol. V. London: Hogarth Press, 1950. Pp. 175-180. (First published in *Int. Z. Psychoanal.,* 1925, *Vol. II.*)

Freud, S. (1926a) *The question of lay analysis.* New York: Norton, 1950. (Translated from *Die Frage der Laienanalyse,* Vienna, 1926.)

Freud, S. (1926b) *Inhibitions, symptoms and anxiety.* London: Hogarth Press, 1936. (Translated from *Hemmung, Symptom und Angst,* Vienna, 1926.)

Freud, S. (1927a) *The future of an illusion.* London: Hogarth Press, 1928. (Translated from *Die Zukunst einer Illusion,* Vienna, 1927.)

Freud, S. (1927b) Postscript to a discussion on lay analysis. In *Collected Papers.* Vol. V. London: Hogarth Press, 1950. Pp. 205-214. (First published in *Int. Z. Psychoanal.,* 1927, *Vol. 13.*)

Freud, S. (1930) *Civilization and its discontents.* London: Hogarth Press, 1930. (Translated from *Das Unbehagen in der Kultur,* Vienna, 1930.)

Freud, S. (1932) Why War? In *Collected Papers.* Vol. V. London: Hogarth Press, 1950. Pp. 273-287. (Written as a letter to Professor Einstein in 1932 and first published in 1933.)

Freud, S. (1933) *New introductory lectures on psycho-analysis.* New York: Norton, 1933. (Translated from *Neue Folge der Vorlesungen zur Einführung in die Psychoanalyse,* Vienna, 1932.)

Freud, S. (1937a) Analysis terminable and interminable. In *Collected Papers.* Vol. V. London: Hogarth Press, 1950. Pp. 316-357. (First published in *Int. Z. Psychoanal.,* 1937, *Vol. 23.*)

Freud, S. (1937b) A disturbance of memory on the Acropolis. In *Collected Papers.* Vol. V. London: Hogarth Press, 1950. Pp. 302-312. (First published in the *Almanach der Psychoanalyse,* 1937, *9.*)

Freud, S. (1938) Some elementary lessons in psychoanalysis. In *Collected Papers.* Vol. V. London: Hogarth Press, 1950. Pp. 376-382. (Written in 1938 and left unfinished.)

Freud, S. (1939) *Moses and monotheism.* New York: A. A. Knopf, 1939. (Parts I and II translated from Moses ein Ägypter and Wenn Moses ein Ägypter war . . . in *Imago,* 1937, BD. XXIII, Heft 1 and 4. Part III written in 1938 and translated into English.)

Freud, S. (1940) *An outline of psychoanalysis.* New York: Norton, 1949. (Translated from Abriss der Psycho-analyse, *Internationale Zeitschrift für Psychoanalyse und Imago,* XXV, 1940.)

Friedlander, Kate *The psycho-analytical approach to juvenile delinquency.* New York: International Universities Press, 1949.

Fromm, E. Die Entwicklung des Christusdogma. *Imago,* 1930, *16,* 305-373.

Fromm, E. *Escape from freedom.* New York: Farrar & Rinehart, 1941.

Green, A. W. Social values and psychotherapy. *J. Pers.,* 1945, *14,* 199-228.

Green, A. W. Sociological analysis of Horney and Fromm. *Amer. J. Sociol.,* 1946, *51,* 533-540.

Hallowell, A. I. Psychic stresses and cultural patterns. *Amer. J. Psychiatry,* 1936, *92,* 1291-1310.

Hallowell, A. I. The Rorschach method as an aid in the study of personalities in primitive societies. *Charact. Pers.,* 1941, *9,* 235-245.

Hallowell, A. I. Acculturation processes and personality changes as indicated by the Rorschach technique. *Rorschach Res. Exch.,* 1942, *6,* 42-50.

Hallowell, A. I. The use of projective techniques in the study of the socio-psychological aspects of acculturation. *J. proj. Tech.,* 1951, *15,* 27-44.

Hartmann, H. Comments on the psychoanalytic theory of instinctual drives. *Psychoanalyt. Quart.,* 1948, *17,* 368-388.

Hartmann, H. Comments on the psychoanalytic theory of the ego. In Anna Freud and others (Eds.), *The psychoanalytic study of the child.* Vol. 5. New York: International Univ. Press, 1950. Pp. 74-96.

Hartmann, H. The mutual influences in the development of ego and id. In Anna Freud and others (Eds.). *The psychoanalytic study of the child,* Vol. 7. New York: International Univ. Press, 1952. Pp. 9-30.

Hartmann, H., & Kris, E. The genetic approach in psychoanalysis. In Anna Freud and others (Eds.), *The psychoanalytic study of the child,* Vol. I. New York: International Univ. Press, 1945. Pp. 11-30.

Hartmann, H., Kris, E., & Loewenstein, R. M. Comments on the formation of psychic structure. In Anna Freud and others (Eds.), *The psychoanalytic study of the child.* Vol. 2. New York: International Univ. Press, 1947. Pp. 11-38.

Hartmann, H., Kris, E., & Loewenstein, R. M. Notes on the theory of aggression. In Anna Freud and others (Eds.), *The psychoanalytic study of the child.* Vols. 3 and 4. New York: International Univ. Press, 1949. Pp. 9-36.

Hendrick, I. Work and the pleasure principle. *Psychoanalyt. Quart.,* 1943, *12,* 311-329.

Henry, J. Some cultural determinants of hostility in Pilagá Indian children. *Amer. J. Orthopsychiat.,* 1940, *10,* 111-119.

Henry, J. Rorschach technique in primitive cultures. *Amer. J. Orthopsychiat.,* 1941, *11,* 230-234.

Henry, J., & Henry, Zunia Doll play of Pilagá Indian children. In C. Kluckhohn, H. A. Murray, & D. M. Schneider (Eds.), *Personality in nature, society, and culture.* New York: Knopf, 1953. Pp. 292-307.

Hilgard, E. R. Experimental approaches to psychoanalysis. In E. Pumpian-Mindlin (Ed.), *Psychoanalysis as science.* Stanford, Calif.: Stanford Univer. Press, 1952. Pp. 3-45.

Horney, Karen *New ways in psychoanalysis.* New York: Norton, 1939.

Jones, E. *Essays in applied psychoanalysis.* London: Hogarth Press, 1923.

Jones, E. Psycho-analysis and anthropology. *J. Royal Anthrop. Inst.,* 1924, *54,* 47-66.

Jones, E. Mother-right and the sexual ignorance of savages. *Int. J. Psychoanal.,* 1925, *6,* 109-130.

Kardiner, A. *The individual and his society.* New York: Columbia University Press, 1939.

Kardiner, A. *The psychological frontiers of society.* New York: Columbia Univer. Press, 1945.

Karpf, Fay B. *American social psychology.* New York: McGraw-Hill, 1932.

Kecskemeti, P., & Leites, N. Some psychological hypotheses on Nazi Germany. I. *J. soc. Psychol.,* 1947, *26,* 141-183.

Kecskemeti, P., & Leites, N. Some psychological hypotheses on Nazi Germany. II. *J. soc. Psychol.,* 1948a, *27,* 91-117.

Kecskemeti, P., & Leites, N. Some psychological hypotheses on Nazi Germany. III. *J. soc. Psychol.,* 1948b, *27,* 241-270.

Kecskemeti, P., & Leites, N. Some psychological hypotheses on Nazi Germany. IV. *J. soc. Psychol.,* 1948c, *28,* 141-164.

Kluckhohn, C. Myths and rituals. *Harvard Theol. Rev.,* 1942, *35,* 45-79.

Kluckhohn, C. Navaho witchcraft. *Papers of the Peabody Museum of Amer. Archaeol. and Ethnol.,* 1944a, *22,* No. 2.

Kluckhohn, C. The influence of psychiatry on anthropology in America during the past one hundred years. In J. K. Hall, G. Zilboorg, & H. A. Bunker (Eds.), *One hundred years of American psychiatry.* New York: Columbia University Press, 1944b. Pp. 589-617.

Kluckhohn, C. The personal document in anthropological science. In L. Gottschalk, C. Kluckhohn, & R. Angell (Eds.), *The use of personal documents in history, anthropology and sociology.* New York: Social Science Research Council. *Bull. 53,* 1945. Pp. 79-173.

Kluckhohn, C. Some aspects of Navaho infancy and early childhood. In G. Roheim (Ed.), *Psychoanalysis and the social sciences.* Vol. I. New York: International Univ. Press, 1947. Pp. 37-86.

Kluckhohn, C., & Morgan, W. Some notes on Navaho dreams. In G. Wilbur & W. Muensterberger (Eds.), *Psychoanalysis and culture.* New York: International Univ. Press, 1951. Pp. 120-132.

Krech, D. Psychological theory and social psychology. In H. Helson (Ed.), *Theoretical foundations of psychology.* New York: Van Nostrand, 1951. Pp. 656-697.

KRECH, D., & CRUTCHFIELD, R. S. *Theory and problems of social psychology.* New York: McGraw-Hill, 1948.

KRIS, E. The significance of Freud's earliest discoveries. *Int. J. Psychoanal.,* 1950, *31,* 108-116.

KRIS, E., & SPEIER, H. *German radio propaganda.* London: Oxford University Press, 1944.

KROEBER, A. L. Totem and taboo: An ethnologic psychoanalysis. *Amer. Anthrop.,* 1920, *22,* 48-55.

KROEBER, A. L. Totem and taboo in retrospect. *Amer. J. Sociol.,* 1939, *45,* 446-451.

KUBIE, L. S. Problems and techniques of psychoanalytic validation and progress. In E. Pumpian-Mindlin (Ed.), *Psychoanalysis as science.* Stanford, Calif.: Stanford University Press, 1952. Pp. 46-124.

LASSWELL, H. D. *Propaganda technique in the world war.* New York: Knopf, 1927.

LASSWELL, H. D. *Psychopathology and politics.* Chicago: University of Chicago Press, 1930.

LASSWELL, H. D. *World politics and personal insecurity.* New York: McGraw-Hill, 1935.

LASSWELL, H. D. *Power and personality.* New York: Norton, 1948.

LEVY, D. Animal psychology in its relation to psychiatry. In F. Alexander & Helen Ross (Eds.), *Dynamic psychiatry.* Chicago: University of Chicago Press, 1952. Pp. 483-507.

LINDZEY, G. An experimental examination of the scapegoat theory of prejudice. *J. abnorm. soc. Psychol.,* 1950, *45,* 296-309.

LINTON, R. (Ed.) *The science of man in the world crisis.* New York: Columbia University Press, 1945.

LOEWENSTEIN, R. M. *Christians and Jews; A psychoanalytic study.* New York: International Univ. Press, 1951.

LOWENFELD, H. Freud's Moses and Bismark. In G. Roheim (Ed.), *Psychoanalysis and the social sciences.* Vol. 2. New York: International Univer. Press, 1950. Pp. 277-292.

LOWENTHAL, L., & GUTERMAN, N. *Prophets of deceit.* New York: Harper, 1949.

MCCLELLAND, D. C., & FRIEDMAN, G. A. A cross-cultural study of the relationship between child-training practices and achievement motivation appearing in folk tales. In G. E. Swanson, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology.* New York: Holt, 1952. Pp. 243-249.

MCDOUGALL, W. *Psycho-analysis and social psychology.* London: Methuen, 1936.

MALINOWSKI, B. *Sex and repression in savage society.* New York: Harcourt, Brace, 1927.

MEAD, MARGARET *Coming of age in Samoa.* New York: Morrow, 1928.

MEAD, MARGARET *Growing up in New Guinea.* New York: Morrow, 1930.

MEAD, MARGARET *Sex and temperament in three primitive societies.* New York: Morrow, 1935.

MEAD, MARGARET Research in contemporary cultures. In H. Guetzkow (Ed.), *Groups, Leadership and men: Research in human relations.* Pittsburgh: Carnegie Press, 1951. Pp. 106-118.

MILLER, D. R., & STINE, MARGARET E. The prediction of social acceptance by means of psychoanalytic concepts. *J. Pers.,* 1951, *20,* 162-174.

MILLER, N. E., & BUGELSKI, R. Minor studies in aggression: II. The influence of frustrations imposed by the in-group on attitudes expressed toward out-groups. *J. Psychol.,* 1948, *25,* 437-442.

MURCHISON, C. (Ed.) *A handbook of social psychology.* Worcester, Mass.: Clark University Press, 1935.

NACHT, S. Discussion of "The mutual influences in the development of ego and id." In Anna Freud and others (Eds.), *The psychoanalytic study of the child.* Vol. 7. New York: International Univ. Press, 1952. Pp. 54-59.

NEWCOMB, T. M. *Social psychology.* New York: Dryden Press, 1950.

NUNBERG, H. The synthetic function of the ego. *Int. J. Psychoanal.,* 1931, *12,* 123-140.

OPLER, M. E. The psychoanalytic treatment of culture. *Psychoanal. Rev.,* 1935, *22,* 138-157.

OPLER, M. E. An interpretation of ambivalence of two American Indian tribes. *J. soc. Psychol.,* 1936, *7,* 82-116.

OPLER, M. E. Further comparative anthropological data bearing on the solution of a psychological problem. *J. soc. Psychol.,* 1938, *9,* 477-483.

PARSONS, ELSIE C. Ceremonial consummation. *Psychoanal. Rev.,* 1915, *2,* 358-359.

PARSONS, T. Certain primary sources and patterns of aggression in the social structure of the western world. *Psychiatry,* 1947, *10,* 167-181.

PARSONS, T. Psychoanalysis and the social structure. *Psychoanal. Quart.,* 1950, *19,* 371-384.

PARSONS, T. Illness and the role of the physician: A sociological perspective. *Amer. J. Orthopsychiat.,* 1951, *21,* 452-460.

PARSONS, T. The superego and the theory of social systems. *Psychiatry,* 1952a, *15,* 15-25.

PARSONS, T. The father symbol: An appraisal in the light of psychoanalytic and sociological theory. Paper read before Chicago Institute of Psychoanalysis, 1952b.

PARSONS, T., & SHILS, E. A. (Eds.) *Toward a general theory of action.* Cambridge, Mass.: Harvard University Press, 1951.

PARSONS, T., BALES, R. F., & SHILS, E. A. *Working papers in the theory of action.* Glencoe, Ill.: Free Press, 1953.

QUEENER, E. L. *Introduction to social psychology.* New York: William Sloane, 1951.

RAPAPORT, D. On the psycho-analytic theory of thinking. *Int. J. Psychoanal.,* 1950, *31,* 161-170.

RAPAPORT, D. The conceptual model of psychoanalysis. *J. Personality,* 1951a, *20,* 56-81.

RAPAPORT, D. The autonomy of the ego. *Bull. Menninger Clin.,* 1951b, *15,* 113-123.

RAPAPORT, D. *Organization and pathology of thought.* New York: Columbia University Press, 1951c.

Reik, T. *Ritual: Psycho-analytic studies.* New York: Norton, 1928.

Riesman, D. *The lonely crowd: A study of the changing American character.* New Haven: Yale University Press, 1950a.

Riesman, D. Authority and liberty in the structure of Freud's thought. *Psychiatry,* 1950b, *13,* 167-187.

Riesman, D. The themes of heroism and weakness in the structure of Freud's thought. *Psychiatry,* 1950c, *13,* 301-315.

Rivers, W. H. R. *Psychology and ethnology.* New York: Harcourt, Brace, 1926.

Roheim, G. Primitive high gods. *Psychoanal. Quart.,* 1934, *3,* 3-133.

Roheim, G. The origin and function of culture. New York: *Nervous and mental disease Monogr.,* 1943, No. 69.

Roheim, G. (Ed.) *Psychoanalysis and the social sciences.* Vol. I. New York: International Univer. Press, 1947.

Roheim, G. *Psychoanalysis and anthropology; Culture, personality and the unconscious.* New York: International Univer. Press, 1950a.

Roheim, G. (Ed.) *Psychoanalysis and the social sciences.* Vol. 2. New York: International Univer. Press, 1950b.

Roheim, G. (Ed.) *Psychoanalysis and the social sciences.* Vol. 3. New York: International Univer. Press, 1951.

Sapir, E. The emergence of the concept of personality in a study of cultures. *J. soc. Psychol.,* 1934, *5,* 408-415.

Scheidlinger, S. *Psychoanalysis and group behavior.* New York: Norton, 1952.

Schneider, L. *The Freudian psychology and Veblen's social theory.* New York: King's Crown Press, 1948.

Sears, R. R. *Survey of objective studies of psychoanalytic concepts.* New York: Social Science Research Council, *Bull. 51,* 1943.

Sears, R. R. Experimental analysis of psychoanalytic phenomena. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* Vol. 1. New York: Ronald Press, 1944. Pp. 306-332.

Seligman, C. G. Anthropological perspective and psychological theory. *J. Royal Anthropological Inst.,* 1932, *62,* 193-228.

Shils, E. A., & Janowitz, M. Cohesion and disintegration in the Wehrmacht in World War II. *Publ. opin. Quart.,* 1948, *12,* 280-315.

Simmel, E. Self-preservation and the death instinct. *Psychoanal. Quart.,* 1944, *13,* 160-185.

Simmel, E. (Ed.) *Anti-semitism: A social disease.* New York: Int. Univ. Press, 1946.

Stanton, A. H., & Perry, S. E. (Ed.) *Personality and political crisis.* Glencoe, Ill.: Free Press, 1951.

Swanson, G. E., Newcomb, T. M., & Hartley, E. L. (Eds.) *Readings in social psychology.* New York: Holt, 1952.

Taylor, W. S. (Ed.) *Readings in abnormal psychology and mental hygiene.* New York: Appleton, 1926.

Tomkins, S. S. (Ed.) *Contemporary psychopathology.* Cambridge, Mass.: Harvard University Press, 1944.

Waelder, R. The principle of multiple function: Observations on over-determination. *Psychoanalyt. Quart.,* 1936, *5,* 45-62.

Waelder, R. *Psychological aspects of war and peace.* Geneva Studies, 1939, *10.* Pp. 1-56.

Waller, W. W. *The family: A dynamic interpretation.* New York: Dryden Press, 1938.

Wilbur, G., & Muensterberger, W. (Eds.) *Psychoanalysis and culture.* New York: International Univer. Press, 1951.

Winch, R. F. *The modern family.* New York: Holt, 1952.

CHAPTER 5

Field Theory in Social Psychology

MORTON DEUTSCH*
New York University

In this chapter, we shall attempt to describe in some detail the theoretical and research contributions of Kurt Lewin and his associates. Since the phrase "field theory" has become closely identified with Lewin's work in social psychology, we shall in a brief introductory section indicate some of the characteristics of field theory in physics as a means of pointing out the rather tenuous connections between "field theory" in physics and in social psychology. In a second section, we shall characterize the main features of Lewin's "way of thinking" about theory in psychology. The third section will deal with the basic concepts of field theory as developed in individual psychology. In a final section we shall discuss "action research" and "group dynamics."

INTRODUCTION

During the latter part of the nineteenth century new and bold ideas were introduced into physics which profoundly altered scientific conceptions of physical reality. The impact of these ideas was not limited to physics. They crumbled the foundations of philosophical viewpoints and served as a model for new conceptions in the biological and social sciences. These ideas — the work of such men as Faraday, Maxwell, and Hertz — are commonly referred to by the term "field theory."

Clerk Maxwell in the Preface to his posthumous book, *Matter and Motion,* characterized the change which field theory brought about in physical conceptions in these words:

> Physical science, which up to the end of the eighteenth century had been fully occupied in forming a conception of natural phenomena as the result of forces acting between one body and another has now fairly entered in the next stage of progress — that in which the energy of a material system is conceived as determined by the configuration, motion of that system, and in which the ideas of configuration, motion, and force are generalized to the utmost extent warranted by their physical definitions.

To put it in other words, Newtonian mechanics had asserted that it was possible to explain physical phenomena by assuming that simple forces were acting between unalterable particles. Hence, if one knew the mass, the velocity, the direction, the location, etc., of material particles one could accurately predict what would occur when a number of them interacted. Every change of motion, according to Newton, was due to a force which arises either through impact — as in the case of a baseball being struck by a bat — or by an attraction exercised mutually by the bodies upon each other, according to his law of gravitation. The gravitational force between bodies (e.g., between the sun and the earth) was assumed to act instantaneously — that is, without time — producing action at a distance. Space and time were treated as the absolute, fixed framework within which the movement of a particle would occur — that is, they were removed from the process itself.

A series of brilliant experiments by Oersted and Faraday on electromagnetic phenomena challenged the notion of simple forces acting instantaneously at a distance on particles in an empty space. These experiments demonstrated that it was useful to think of regions or fields through which electromagnetic forces were spread and to assume that these forces existed even when there were no material particles present to which the forces were applied. It was found that a knowledge of the properties of the electromagnetic field was sufficient to explain electromagnetic phenomena and that, given this information, knowledge about the source of the electromagnetic field was unnecessary. In other words, the electromagnetic field had a definite reality which was not in any way the end result of distinct, individual particles or sources. As Einstein and Infeld (1938, p. 259) have pointed out: "It needed great scientific imagination to realize that it was not the charges nor the particles but the field in the space between the charges and particles which is essential for the description of physical phenomena."

---

* I wish to express my indebtedness to a number of colleagues for their critical reading of a first draft: Drs. Kurt Back, David Emery, Harold Gerard, John Harding, Philip Lichtenberg, and Marie Jahoda. In addition, this chapter is indebted to an unpublished manuscript which was written with the collaboration of Dr. Emery.

Clerk Maxwell systematized and gave mathematical form to these experimental findings in his field equations. These equations, unlike the equations of Newtonian physics, are *structural* equations which rule out action at a distance. Einstein and Infeld (pp. 152-153) have described the difference between mechanical laws and field laws in these terms:

> All space is the scene of these laws and not, as for mechanical laws, only points in which matter or charges are present . . . Maxwell's equations enable us to follow the history of the field, just as the mechanical equations enabled us to follow the history of material particles . . . However, they do not, as in Newton's laws, connect two widely separated events; they do not connect the happenings *here* with the conditions *there*. The field *here* and *now* depends on the field in the *immediate neighborhood* at a time *just past* . . . They allow us to increase our knowledge of the field by small steps. In Newton's theory, on the contrary, only big steps connecting distant events are permissible.

Einstein's theory of gravitation, formulated at the beginning of the twentieth century, routed the concept of action at a distance from its last remaining stronghold in the physical sciences. The notion of a space filled with material ether vanished from physics and was replaced by a conception of space as a definitely distributed system of gravitational and electromagnetic forces and stresses. In the field conception the distribution of forces in a given environment determines what an object with certain properties will do in that environment. Thus, when we know the object and observe what it does in a specified environment we can deduce the properties of the field in that environment.

The field concept in physics has served as a stimulating analogy for scientific workers in other areas. Essentially, "field theorists" in the nonphysical sciences have attempted to consider the phenomena they investigate as occurring in a "field" — i.e., as part of a totality of coexisting facts which are conceived of as mutually interdependent. In other words, "field theory" as it is employed in the social and biological sciences does not refer to theories about physical (e.g., gravitational or electromagnetic) phenomena; rather it refers to a "method of analyzing causal relations and of building scientific constructs" (Lewin, 1951b, p. 45, or 1943a). This method, in a manner analogous to that of field theory in physics, assumes that the properties of any event are determined by its relations to the system of events of which it is a component and that "changes here and now depend on changes in the immediate neighborhood at a time just past."

The term "field theory" in psychology has been applied primarily to the work of the Gestalt psychologists, and it has been employed to characterize especially the work of Kurt Lewin and his students. However, in view of our previous discussion of the meaning of "field theory," it should be evident that there are many possible approaches to theorizing in the social sciences which could be "field theoretical" in spirit. In the domain of social psychology, the theoretical writings of Asch, Krech and Crutchfield, Festinger, Sullivan, Gardner Murphy, Newcomb, Coutu, Brown, G. H. Mead, Parsons, Cottrell, Barnard, Homans, Simon, Heider, Barker, Tolman, Wright, the staffs of the Research Center for Group Dynamics and of The Tavistock Institute, and many other equally prominent theorists are more or less "field-theoretical" in their approach.

THE METATHEORY

Kurt Lewin and his students have made many important substantive contributions to the field of social psychology. It may well be, as some critics have asserted, that Lewin's brilliant experimental studies are not necessary implications of his theoretical statements or of the field theoretical approach. Nevertheless, anyone who has had the opportunity to watch Lewin at work or who has attended the annual meetings of the "Topological Psychologists" (a group of people sympathetic to Lewin's approach) has very likely sensed a characteristic "way of thinking." It can be sensed in the method of formulating questions for psychological research, in the emphases and neglects, in the manner in which ideas are linked together, and in the insights achieved.

We shall try to describe some of the main features of the way of thinking which characterizes the "field-theoretical" approach to psychology. Many of the features of this way of thinking are not unique to Lewin's approach. The wide impact that Lewin and Gestalt psychologists have already had on theorizing in psychology as well as an increasing interest in the nature of theory construction among all schools of psychology have led to a considerable overlap in their basic viewpoints.

A good proportion of Lewin's writings are devoted to the metatheory of psychology. We shall not attempt to quote his writings in any detail but rather shall attempt to present the essentials of the way of thinking which characterized his approach. The reader, however, may find it useful to consult his relevant writings in *Field Theory in Social Science, A Dynamic Theory of Personality, Principles of Topological Psychology,* and *The Conceptual Representation and the Measurement of Psychological Forces.*

We shall discuss the characteristics of Lewin's approach to theory building under the following headlines: A. The Psychological Approach, B. An Emphasis on the Total Situation, C. Systematic Rather than Historical Causation, D. The Constructive Rather than Classificatory Approach, E. The Dynamic Approach, F. Mathematics and Field Theory.

THE PSYCHOLOGICAL APPROACH

Despite Lewin's use of terms which have the flavor of physics — e.g., *tension, force, fluidity* — his theoretical approach is consistently *psychological.* His use of concepts like "tension" and "force" is not an attempt to derive psychological from physical concepts but rather reflects his view that what he termed the "logic of dynamics" might be similar in the various sciences. He takes the stand that psychological phenomena must be explained in psychological terms just as physical phenomena must be explained in physical terms.

Let us briefly attempt to characterize what is meant by an explanation of psychological phenomena in psychological terms. [Our discussion is much indebted to Heider (1927), Wright and Barker (1950), and Tolman (1932) and relates to the molar-molecular distinction which Littman and Rosen (1950) have thoroughly reviewed.] First of all, this approach assumes that it is scientifically respectable to describe the phenomena in which psychologists are interested in terms that are not reducible or equivalent to the terms of physics or physiology. Secondly, it postulates that psychology is concerned with the individual's behavioral transactions with his internal and external environment. Behavior is a transaction of an individual, not merely of some part of the body. As Wright and Barker (1950, p. 79) point out, the individual:

> . . . does not sweat or salivate, nor does he often bend his knees in walking, manipulate his tongue in talking, move his eyeballs in reading, or bend at the waist in sitting down. *He* walks, talks, reads, or sits down, leaving his glandular and motor apparatus to take care of the sweating, salivating, bending, manipulating and all such molecular units of behavior which, as molecular, are lost to the person in what he actually does.

The behavioral transactions of an individual with his environment are mediated by physical and physiological processes; in other words, these processes are the mechanisms or tools for behavior, not behavior itself. It would make no more sense to describe the latter as behavior than to describe the movement of a steam shovel as behavior when a construction worker is employing it to dig a foundation for a building. This is not to assert that mediating tool processes may not at one time or another be the object of behavior — this is particularly likely to be the case when the individual is learning how to use the given tool; e.g., as the baby learns to walk, his behavior is directed at movements of his legs.

The psychological explanation of behavior assumes that all behaviors have directional characteristics. Hence, it is concerned with the purposes which underlie behavior and the goals toward or away from which behavior is directed. There is little direct interest in tools or mediating processes *per se* apart from the interest in how they are learned or acquired and in how desirable and effective they are considered to be by the individual.

An emphasis on the psychological explanation of psychological events in contrast to an explanation in terms of physical or neurophysiological processes has several important consequences. The scientist who works with the assumption that psychological phenomena are realities, even if they cannot be expressed in physical terms or located in physical space, is more apt to investigate them. Recognizing that people in everyday affairs use such concepts as "hope," "desire," "action," "ability," "ought," etc., to characterize others as well as themselves (and acknowledging that these phenomena have reality in their own right), he is more likely to feel that the concepts of this everyday psychology are legitimate and important objects for investigation by a scientific psychology (Heider, 1953). That is, he will want to know under what conditions a person with given characteristics will say he is "hopeful" and under what conditions this person will say that another person is "hopeful." In other words, he is more apt to concern himself with the psychological phenomena which occur in everyday life, including the concepts of everyday psychology, than is the scientist who feels that he is not being scientific unless he can either say "hope" is really a specified pattern of neural excitation in a certain region of the brain or that "hope" is really a certain pattern of bodily movements which can be expressed in physical terms. This is not to assert that the psychologist who employs a psychological approach in theorizing will be content to stop at the level of description of the phenomenal properties of psychological events. He will want to know how "hope" can be characterized in terms of the constructs of a psychological theory. Lewin, in his writings, continuously stresses that the understanding of causal relationships requires one to proceed from the more accessible, phenomenal properties of psychological events to their underlying dynamic or genotypical properties.

The emphasis on a psychological explanation of psychological events leads to an avoidance of "achievement concepts" (i.e., of concepts which define behavior in terms of what it accomplishes) since, as Lewin (in Rapaport, 1950, p. 90) writes, it is impossible "to subsume, under unitary psychological laws, processes defined by achievements." More generally, Lewin objects to the use of simple external criteria to identify or define a psychological state or process. In effect, he asserts that it is poor psychological theorizing to conclude that a person who is exposed to a given external event (e.g., of being "rewarded") will necessarily interpret or experience the event as it was intended. Similarly, it is poor theorizing to assume that all of the effects of a person's behavior (e.g., resulting in another person "feeling insulted") are psychologically determined. Lewin's point is that to develop psychological laws we must deal with psychological processes — the perception of the external event, not the event by itself; behavior, and not its effects by themselves. Thus, we ought to be concerned with the psychological impact of a "reward," e.g., whether it is perceived as a "bribe" or as a "sign of accomplishment," rather than simply with the external situation of reward. In essence, one has to deal with what *exists psychologically*, what is real for the person being studied. If a stimulus object does not exist for an individual, consciously or unconsciously, the object cannot be thought of as having any psychological effects; similarly, if the behavior effect was not consciously or unconsciously motivated, it cannot be thought of as being psychologically determined.

Emphasis on the study of the person and the psychological environment as it exists for the person has led critics (e.g., Brunswik, 1943) to characterize Lewin's approach as "postperceptual and prebehavioral." Lewin's lack of detailed discussion of perception and of action provide a justification for this criticism. Nevertheless, Lewin denied this, asserting that perception and action are, in his view, legitimate problems of psychology. He wrote (Lewin, 1951b, p. 57, or 1943a):

> This view is a necessary consequence of the field-theoretical approach according to which the boundary conditions of a field are essential characteristics of the field. For instance, processes of perception which should be related to the boundary zone depend partly on the state of the inner part of the psychological field, i.e., upon the character of the person, his motivation, his cognitive structure, his way of perceiving, etc., and partly on the "stimulus distribution" on the retina or other receptors as enforced by physical processes outside the organism.

Lewin was fully aware that the scientist could not depend upon intuition in order to apprehend what was psychologically real for another person. Rather, he had to infer or construct this reality from observable data, such as the course and consequences of the person's actions and the situations in which action was occurring. Further, he was aware that "the investigation and representation of each single case is an infinite task in itself and one whose solution would presuppose a full knowledge of the laws of psychology" (Lewin, 1936a, p. 17). Nevertheless, his interests and his views as to a fruitful strategy for psychology led him to a relative neglect of problems of perception and action in favor of an emphasis on motivational problems. As a consequence, he used a rather "make-do" methodology for identifying the specific properties of a given individual's psychological field or life space. The tool most commonly employed in experimental studies was simply that of verbal report: the subject described in his own terms what a situation meant to him or what he intended. Despite Lewin's awareness of the liabilities in employing verbal report, he was not blinded to its assets when studying capable human subjects.

Lewin's use of verbal report as a "make-do" methodology and his emphasis on describing the psychological environment in terms of "what exists for the person" rather than in "objective *physicalistic*" terms has led some critics (e.g., Spence, 1944) to mistakenly characterize Lewin's theoretical approach as "phenomenological." Lewin, first of all, never equated "what exists for the person" with "awareness," "consciousness," or "ability of the person to describe verbally"; he realized that many important psychological influences operate unconsciously. Thus, in the sense of being only concerned with what exists in awareness, Lewin was not a phenomenologist. Secondly, Lewin realized, as do all sophisticated scientists, that the concepts and methods of psychology had to be intersubjective and, in this sense, psychological processes in another person could only be apprehended from externally observable data. Hence, in the sense that Lewin never assumed that verbal report or phenomenological descriptions were simple portrayals of the "raw truth," he was also not a phenomenologist. On the other hand, he was interested in what people are aware of and he was sensitive to the insights into psychological processes which could be obtained from subjective reports.

Although Lewin believed that the employment of achievement concepts would prevent the development of unitary, psychological laws, he recognized that any long-range or detailed prediction of individual behavior required a knowledge of nonpsychological facts — a knowledge of facts which, though not affecting the individual at a given time, will come to affect him at a later time. For example, if one wishes to predict what a particular individual is going to do during a Sunday drive into the country one will have to be able to predict whether or not his automobile will get a punctured tire, whether or not there will be a traffic jam, etc., as well as what he seeks in his excursion. The specific prediction of what a person will do or what he will become at some future time requires a knowledge of the future situations to which he will be exposed. The task of discovering what part of the physical or social world will determine during a given period the "boundary zone" of the life space has been called "psychological ecology" by Lewin. He and many of his students have devoted much of their research to studies of the psychological ecology of different social settings. We shall discuss psychological ecology in greater detail in a subsequent section of this chapter.

AN EMPHASIS ON THE TOTAL SITUATION

The most fundamental construct for Lewin is that of the psychological "field" or life space. All psychological events (thinking, acting, dreaming, hoping, etc.) are conceived to be a function of the life space which consists of the person and the environment viewed as *one* constellation of interdependent factors. That is, all psychological events are conceived to be determined, not by isolated properties of the person or his environment, but by the mutual relations among the totality of coexisting facts which comprise the life space, facts which derive from the momentary condition of the individual and the structure of his environment. Individual psychological processes are, in Lewin's words, *"always to be derived from the relation of the concrete individual to the concrete situation,* and, so far as internal forces are concerned, from the mutual relations of the various functional systems that make up the individual" (Lewin, 1935b, p. 41).

The emphasis on the interrelatedness of the person and the environment was one of Lewin's major contributions to psychological theorizing. Until recently, much of psychology has been dominated by what he has termed an "Aristotelian" mode of thinking (Lewin, 1935b, Chap. 1). In such thinking, psychological events are determined by the characteristics of the individual — "instincts," "heredity," "intelligence," "needs," "habits," etc. — relatively independent of the situation. Since Lewin it has become increasingly evident that it is meaningless to speak of behavior without reference to both the person and his environment. Statements such as the following, which ignore the situation, are viewed as unacceptable by an ever-increasing number of psychologists: "the reason he is psychotic is because of his heredity"; "the rigidity of his problem-solving behavior is due to his ethnocentrism"; "he became the leader of his group because of his personality"; "her emotional outburst was due to her hysteria"; "friends work better together than do strangers"; "a highly cohesive group will be more productive than a less cohesive group."

Lewin's emphasis on the "relation of the concrete individual to the concrete situation" leads the scientific investigator to a more explicit realization that understanding behavior requires not only a knowledge of the person (e.g., his past experiences, his present attitudes, and his capabilities) but also a knowledge of his immediate situation. This is true for understanding everyday behavior and also for understanding the behavior of subjects in carefully controlled experiments or diagnostic testing situations. In these latter instances, interpretation of the responses of a subject requires a knowledge of the psychological significance of the total experimental or testing situation. To cite a simple example, the responses of subjects in a locked room to smoke pouring through the cracks in the floor and to the "shouts" of fire from outside will be determined largely by whether they view the smoke to be part of an experimental hoax or to be part of a real fire (French, 1944).

The emphasis on the total situation has consequences for research methodology. Instead of picking out one or another isolated element within a situation, field theory finds it advantageous, as a rule, to start with a characterization of the situation as a whole. After this first approximation, the various aspects of the situation undergo a more and more detailed analysis. The procedure:

> . . . is not to add disconnected items but to make the original structure more specific and differentiated. This method therefore proceeds by steps from the more general to the particular and thereby avoids the danger of a "wrong simplification" by abstraction. (Lewin, 1936a, p. 17)

In essence, the emphasis in research guided by field theory is to study psychological events in terms of the units of which they are a part — to study them *in their relations of interdependence* rather than as isolated abstractions torn from their relational significance. This emphasis leads to a methodological interest in studying "pure" cases (that is, cases where accidental or irrelevant variables are eliminated as far as possible) in sufficient detail to understand psychological events in terms of the totality of their mutual interrelations. Contrasted with this approach would be the study of the relationships between variables — e.g., intelligence and school achievement — by taking a large number of cases, none of which are studied in any detail from the point of view of their concrete relations to their specific situations. The assumption sometimes underlying this latter approach is that the relationship between variables such as intelligence and school achievement is independent of the other properties of the situation or of the person. Field theorists, on the other hand, would assert that the relationship between any two variables of this sort is almost inevitably influenced by the total situation, and thus they would conclude that their relationship should be established only after the properties of the total situation have been sufficiently characterized.

SYSTEMATIC VERSUS HISTORICAL CONCEPTS OF CAUSATION

In the introductory section of this chapter, it was pointed out that field theory in physics destroyed the notion of "action at a distance" and replaced it by structural equations in which "the field here and now depends on the field in the immediate neighborhood at a time just past." Lewin, in a similar vein, has emphasized that psychological events (i.e., changes in the life space) have to be explained in terms of the properties of the field which exist *at the time* when the events occur. He points out that past events can only have a position in the historical causal chains whose interweaving creates the present situation; they cannot directly influence present events.

This is not to deny the significance of the past in indirectly affecting behavior. However, even though a past event can create a certain condition which carries over into the present, it is, nevertheless, the *present* condition that is influential in the present. Strictly considered, linking behavior with a past event is an extremely difficult undertaking; it presupposes that one knows sufficiently how the past event affected the psychological field at that time, and whether or not in the meantime other events have again modified the field. While the notion that only the present is influential in the present is an obvious concept, one still can find in abnormal psychology textbooks accounts of how an event in childhood "caused" a neurosis. A more adequate explanation would have to describe the "neurotic trends" resulting from the childhood event, the vicissitudes of these trends during the period between the event and the neurosis, and the interactive nature of the present situation and neurotic trends which "result" in the neurotic symptoms.

While Lewin emphasized the principle of contemporaneity as one of the basic characteristics of field theory, he realized that the determination of the meaning of any psychological event requires the taking into account of an appropriate time period. His use of the term "contemporaneous" or the phrase "situation at a given time" actually does not refer to an instant without time extension, but to a certain time period. Time differentials in psychology — i.e., what is viewed as "contemporaneous" — must be conceived in terms of the duration of meaningful psychological units; there is no reason to make the *a priori* assumption that psychological space-time is infinitely or arbitrarily divisible. That is, psychological units have to be thought of as having an extension in regard to both field and time dimensions, and unless the duration of the time-period is appropriate to the unit meaningful psychological units will not be observable.

Lewin's emphasis on employing units of sufficient time extension has methodological implications. He writes that "the first prerequisite of a successful observation in any science is a definite understanding about *what size of unit* one is going to observe at a given time" (Lewin, 1951b, p. 157, or 1943b). He points out that in social psychology we have often misinterpreted the scientific requirements of analysis and have tried to observe as small units as possible and, in so doing, we have frequently torn the observed units from their context. The fallacy in such an approach is that there is frequently no way to distinguish among different possibilities of classifying an action if the observation lasts only a few seconds. Thus, if two persons *A* and *B* are running one behind the other, it may mean that either *A* is leading *B* or that *B* is chasing *A*. Only by observation which is sufficiently comprehensive to include both *A* and *B* and to extend through a period sufficiently long does the meaning of the activity become clear. (For a further discussion of systematic versus historical causation see Chein, 1947.)

THE "CONSTRUCTIVE" RATHER THAN "CLASSIFICATORY" APPROACH

Lewin's orientation to science was much influenced by the work of Ernst Cassirer, a German philosopher who was concerned with the nature of concepts in mathematics and the sciences. In his book, *Substance and Function and Einstein's Theory of Relativity,* Cassirer (1923) emphasized the opposition between two modes of theorizing which are distinguished by the different value which each places upon *thing concepts* and *relation concepts.* The classificatory approach, which is characterized by thing concepts, assumes that concepts are derived by abstractions from particular objects to an abstract object which embraces the essence of the particular object to which it refers. Relations or connections, as a rule, are considered among the nonessential properties of a concept, and thus as capable of being left out of its definition without fallacy. In the classificatory approach, as we go to "higher" and "higher" concepts (e.g., from "cow" to "animal" to "living object," etc.), the conceptual pyramid which we form in this way reaches its summation in the abstract representation of "something" — the all-inclusive "entity" to which every possible content is relevant, but which at the same time is totally devoid of specific meaning.

The constructive approach to theorizing stresses relational concepts; it points out that even the process of abstraction presupposes a relational principle by which objects are identified as having a common characteristic. In this approach, the meaning of any concept is determined by its relation to other concepts in the system of concepts of which it is a part. The reality of specific phenomena are derivable or constructible from the relevant constructive elements or constructs.

The ideal of a scientific concept held by the constructive approach is not an abstraction which disregards the peculiarities and particularities of the specific objects to which it is relevant but rather a concept which seeks to show the necessity of the occurrence and connection of just these particularities. In other words, according to the constructive approach, with adequate scientific concepts, there is no antithesis between general laws and the individual case. By introducing parameters which are appropriate to the individual case we can derive individual behavior from a formula which is applicable to all cases.

Lewin characterizes his approach to the development of psychological theory as the constructive method, and he attacks the attempt to develop concepts by a process of abstracting from phenotypical or historic-geographic data (i.e., data in which the similarities are defined in terms of external rather than psychological characteristics). He writes:

> If one "abstracts from individual differences," there is no logical way back from these generalizations to the individual case. Such a generalization leads from individual children to children of a certain age or certain economic level and from there to children of all ages and of all economic levels; it leads from the psychopathic individual . . . to the general category, "abnormal person." However, there is no logical way back from the concept "child" or "abnormal person" to the individual case. What is the value of general concepts if they do not permit prediction for the individual case? (Lewin, 1951b, p. 61, or 1942)

Phenotypical data must be transformed into the language of constructs ("genotypes," to use Lewin's term) before one can hope to develop general psychological laws which will be applicable to individual cases. The development of general psychological laws, however, not only requires the coordination of data to constructs by definite rules of correspondence and explicit operational procedures, but it also presupposes that the linkages or relations among constructs are guided by an awareness of the conceptual dimensions of the constructs. Lewin (1951b, p. 37, or 1944) points out that "(1) Only those entities which have the same conceptual dimension can be compared as to their magnitude. (2) Everything which has the same conceptual dimensions can be compared quantitatively; its magnitude can be measured, in principle, with the same units of measurement."

While Lewin does not discuss "conceptual dimensions" at any length, he points out several major distinctions, which we shall discuss more fully as we discuss his specific concepts. Briefly, he distinguishes between structural concepts (e.g., position, cognitive structure) and dynamic concepts (e.g., force, tension); between concepts which refer to different regions of the time dimension; between concepts which refer to different levels of reality; between concepts referring to "potential" and concepts referring to "actual"; etc. For example, "position" is a spatial relation of regions, while "locomotion" refers to a "relation of positions at different times." "Hope" refers to a relation between the structure of the reality level and of the wish level of the psychological future. "Guilt" refers to the relation between the structure of the reality and the wish level of the psychological past.

Lewin's stress on clarification of the conceptual dimensions of psychological constructs reflected his feeling that the emphasis on "operational definitions" of scientific concepts was being misconstrued in psychology. He felt that some psychologists naïvely assumed that the meaning of a concept is exhausted by its operational definition — e.g., "intelligence is what is measured by intelligence tests." Lewin, on the other hand, emphasized that completely adequate scientific concepts not only required explicit operational definitions which would provide linkages to observable facts but also conceptual definitions which established the linkage between any given concept and other concepts in the system of concepts (or theory). One might go a little further and state that the possibility of an unequivocal and consistently repeatable operational definition of any psychological concept is dependent upon the state of psychological theory. That is, one cannot give a precise statement of what is being done (i.e., a statement of all the relevant variables which interact with one's measurement procedures and the variable being measured) as one measures "intelligence," for example, unless one has a theory which enables specification of the variables which may influence either "intelligence" in the situation of measurement or the testing procedure by which "intelligence" is measured.

THE DYNAMIC APPROACH

Lewin frequently uses the phrase "dynamic concepts" as an equivalent for such terms as "constructs," "genotypes," "conditional-genetic concepts," and "intervening concepts." In such a usage, any approach to psychology which employs constructs to systematize and interpret observable data may be called "dynamic." One could say that practically all major approaches to psychology are, in this sense, dynamic. However, a more limited use of the term "dynamic" is common. In this more limited usage, the common denominator is the conception of living things as systems which tend to maintain a *dynamic equilibrium* in relation to their environments. That is, the individual is conceived of as a system whose components maintain a relative flexibility of interrelations so that variation in one aspect of the system may induce compensating changes in other aspects, thus maintaining the integrity of the system. The notion of dynamic or organismic equilibrium stresses the fact that the "whole" remains the same only because of a patterning or organization which persists in the midst of change. Contrasted with the model of dynamic equilibrium is that of mechanism. The model of mechanism

> . . . implied the notion of a whole which was completely equal to the sum of its parts; which could be run in reverse; and which would behave in exactly identical fashion no matter how often those parts were disassembled and put together again, and irrespective of the sequence in which the disassembling or reassembling would take place. It implied consequently that the parts were never significantly modified by each other, nor by their own past, and that each part once placed into its appropriate position with its appropriate momentum, would stay exactly there and continue to fulfill its completely and uniquely determined function. (K. Deutsch, 1951, p. 234)

Arising out of the conception of the individual as an "equilibrium-maintaining system" is an interest in the processes by which equilibrium is restored when it is disturbed. In psychology, this interest is reflected in an emphasis on motivational processes — e.g., the arousal of need-tensions, the setting of goals, goal-directed action, the release of tension. Much of the research of Lewin and of his coworkers was directed toward an understanding of motivational processes. Lewin's theoretical approach is, in the sense of the preceding sentence, legitimately grouped with the approaches of McDougall and Freud as a motivational or "dynamic" psychology.

All theories in psychology which use the concept of equilibrium include more or less conspicuously the thesis that behavior or any other kind of psychological change is caused by directed entities. Hull's concept of "excitatory tendency," McDougall's "propensities," Freud's "libido," Tolman's "theory of purposive behaviorism," etc., all imply directed entities. In Lewin's system these directed entities are called "psychological forces." In field theory, the vectorial nature of "psychological forces" is explicitly recognized and an attempt is made to develop a mathematics which would adequately handle for psychology the concepts of magnitude and direction implicit in the concept of vector.

MATHEMATICS AND FIELD THEORY

Lewin stressed the need for clear understanding of the formal properties of scientific constructs, and he insisted that the determinants of behavior would have to be represented in mathematical terms if psychology were to become a rigorous discipline. He devoted great effort in attempts to apply and to formulate geometrical concepts which would be of value in psychology. His most elaborate effort in this connection was his formulation of a geometry which he termed "hodological space" (Lewin, 1938).

Although it cannot be said that "hodological space" was well developed from the mathematical point of view, it represents an extremely ingenious attempt to develop a geometry whose basic spatial concepts could be integrated with dynamic concepts. In hodological space, the distinguished path which is used in determining the direction between any two points is determined by the attractiveness rather than the shortness of the path. Direction is influenced by such factors as the degree of differentiation of the space into subregions, the relative prominence of wholes versus parts, the properties of the field at large. Using the concepts of hodological space to define psychological direction seems particularly helpful in giving insight into behavior in conflict situations, in situations which are cognitively unclear, in situations involving roundabout routes, in understanding the difference between escape (or withdrawal) and goal-seeking (or approach) behavior, etc. It permits an answer to the puzzling necessity to ascribe different psychological directions to behaviors in the same physical direction if the goals of these behaviors are different.

Although Lewin was constantly directing his energies to the development of a more adequate formalistic system, he felt that an enthusiasm for formal systems might lead to a substitution of mere verbalisms for empirically descriptive theories. Formalization and mathematization, if prematurely done, may lead us to the building of a logical superhighway which turns out to be a "dead end leading nowhere." A young science, if it is wise,

> . . . follows the same procedure used in the exploration of the resources of a new land: small paths are pushed out through the unknown; with simple and primitive instruments, measurements are made; much is left to assumption and to lucky intuition. Slowly certain paths are widened; guess and luck are gradually replaced by experience and systematic exploration with more elaborate instruments. Finally highways are built over which the streamlined vehicles of a highly mechanized logic, fast and efficient, can reach every important point on fixed tracks. (Lewin, 1951b, p. 3, or 1940)

BASIC CONCEPTS OF FIELD THEORY AS DEVELOPED IN INDIVIDUAL PSYCHOLOGY

Although much of Lewin's writing is concerned with the metatheory of psychology, he also has made a major effort to develop a specific psychological theory. In this section, we shall outline some of the major concepts employed in his theory — particularly the concepts which were developed in his study of individual psychology — and shall attempt to indicate how these concepts were employed in research. It is well to realize at the outset that Lewin's theoretical system was never fully developed nor was it formalized into a hypothetico-deductive system. As a consequence, there is a vagueness and lack of precision about many of the concepts; a vagueness which reflects Lewin's feeling that theory had to be constructed by a method of approximation.

We shall divide our discussion into four main parts: (A) The life space, (B) Structural concepts, (C) Dynamic concepts, (D) Concepts dealing with change in the psychological environment.

THE LIFE SPACE

As we have pointed out in an earlier section, Lewin employs the term "life space" (or "psychological field" or "total situation") to refer to the manifold of coexisting facts which determine the behavior of an individual at a certain moment. Behavior, in other words, is a function of the life space: $B = f\ (LS)$. The life space is in turn a product of the interaction between the person (P) and his environment (E). In symbolic expression, $B = f\ (LS) = f\ (P, E)$. Unfortunately, Lewin used each of his crucial terms — "environment," "person," and "behavior" in several different ways.

Environment. The term "environment" is often used to refer to the objective environment or stimulus situation (i.e., the objective situation which is presented to the individual at a given moment: the situation which acts upon the individual's perceptual apparatus and upon which his motor apparatus acts). It is this meaning of the term "environment" which is ap-

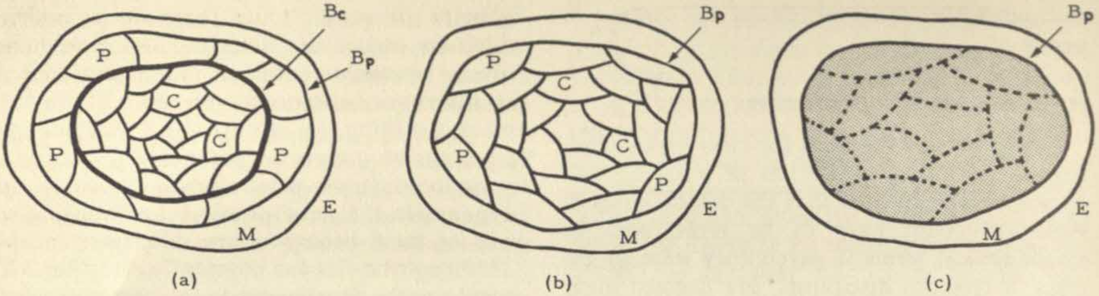


Fig. 1. THE REPRESENTATION OF THE PERSON. Relations among various strata of the person in different circumstances. (a) The person in an easy situation: the peripheral parts P of the inner personal region I are easily accessible from outside, E; the more central parts C are less accessible; the inner personal region I influences the motor region M relatively freely. (b) The person under stress, in state of self-control: the peripheral parts P of the inner personal region I are less accessible than in (a); peripheral and central parts (C and P) are more closely connected; communication between I and M is less free. (c) The person under very high tension: unification (primitivation, "regression") of the inner personal region I.

Symbols: M, motor-perceptual region. I, inner personal region. P, peripheral parts of I. C, central parts of I. E, environment. B_C, dynamic wall between C and P. B_P, dynamic wall between I and M.

(Adapted from K. Lewin, *Principles of Topological Psychology*, p. 181.)

propriate to the statement that the life space is "a product of the interaction between the person and his environment." Similarly, as Leeper (1943) has pointed out, this meaning is appropriate to such statements as the following:

> The motor and perceptual region has the position of a boundary zone between the inner personal regions and the environment . . . Two groups of facts stand in favor of such a representation.
>
> 1. Needs or other states of the inner personal regions can influence the environment only by way of a bodily expression or a bodily action, that is, by way of a region which one can call the motor region.
>
> 2. We find an intermediate region again when we consider the influence in the opposite direction, namely, psychological changes of the inner personal region resulting from changes of the environment. This intermediate region corresponds to the perceptual system in the broadest sense of the word, that is, to hearing, seeing, etc. (Lewin, 1936a, pp. 177-179)

Lewin also uses the term "environment" to refer to the "psychological environment," which is conceived to be the environment as it "exists" for the individual. The psychological environment is a part of the life space and, hence, its properties are determined not only by the characteristics of the objective environment but also by the characteristics of the person. The fact that the psychological environment is an interactive product is most evident in instances where the psychological environment is not veridical to the objective environment — e.g., consider any of the common perceptual distortions under intense need. In addition, a little thought will reveal that regions of the psychological environment often have certain characteristics which do not exist in the objective environment independently of its relation to a given person — e.g., the attractiveness or repulsiveness of certain objects.

Person. Lewin employed the term "person" in three ways. First of all, he used the term to refer to those properties of the individual (e.g., his needs, his beliefs and values, his perceptual and motive systems) which in interaction amongst themselves and with the objective environment produce the life space. One might say that here Lewin is referring to the individual or organism considered "psychologically" rather than physiologically. One could say that "person" in this sense is the life space at a time just past which, as a consequence of its interaction with a subsequent objective environment, has produced the present life space. In a second usage, Lewin also employs the term "person" as being equivalent to "life space." His representation of the "person" (see Fig. 1) and his representation of the "life space" (Fig. 2) may be viewed as different ways of representing essentially the same psychological facts. Figure 2 is more convenient for representing changes in the psychological environment, and Fig. 1 for representing changes in tension systems.

Lewin also used the term "person" to refer to the "person in the life space." The person in the life space or "the behaving self," to borrow a more expressive term employed by Tolman (in Parsons and Shils, 1951), is the individual as related to the other entities in his life space. Just as the "psychological environment" may be thought of as the "perceived environment," so the "behaving self" may be

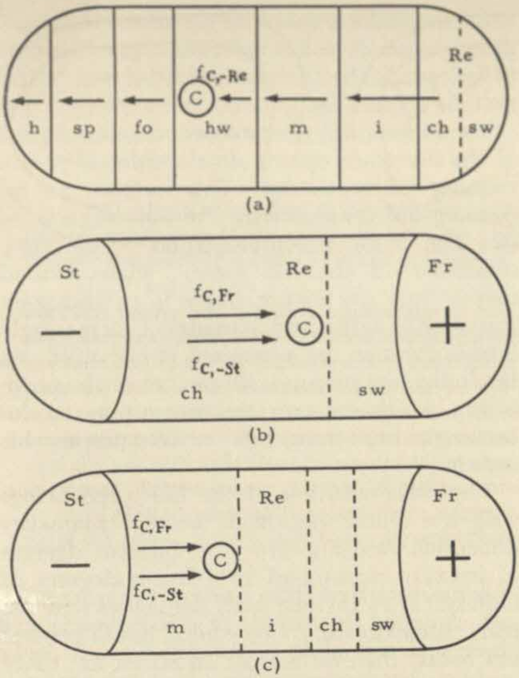


FIG. 2. THE REPRESENTATION OF THE LIFE SPACE. FINITELY STRUCTURED FORCE FIELD CONSISTING OF MACROSCOPIC REGIONS. (a) Eating situation in case of disliked food. *C*, child; *Re*, real eating; *h*, putting hand on table; *sp*, taking spoon; *fo*, putting food on spoon; *hw*, bringing spoon halfway to mouth; *m*, bringing spoon to mouth; *i*, taking food into mouth; *ch*, chewing; *sw*, swallowing; eating has a negative valence $[Va(Re) < 0]$; the force away from real eating, $f_{C,-Re}$, increases stepwise with the decrease of distance between *C* and *Re*.

(b) Change of direction of forces after the child started real eating. *C*, child; *Re*, real eating; *St*, struggle with adults; *Fr*, freedom; $f_{C,Fr}$, force in the direction to freedom; $f_{C,-St}$, force away from struggle.

(c) In a later stage of "learning" to eat the disagreeable food, the situation might be restructured so that bringing spoon to mouth *(m)* is now seen as a subpart of the region of real eating *(Re)*.

(Adapted from K. Lewin, *The Conceptual Representations and Measurement of Psychological Forces*, p. 117.)

thought of as the individual's perception of his relations to the environment he perceives. The psychological environment and his behaving self are interdependent components of the life space. Most observable behavior can be thought of as paralleling changes of position of the behaving self from one activity region to another in the life space. It should be stressed that the properties of the behaving self are completely expressed in a statement of its relations and relative position to the other objects in the individual's life space. Thus, there are no internal relations within the behaving self, and hence it may be considered as an undifferentiated region for representational purposes.

There is some evidence to indicate that Lewin would have disagreed with our view that he employs the term "person" in the different meanings described above. For example, in his *Principles of Topological Psychology* he has a chapter entitled "The Person as a Differentiated Region in the Life Space" and in this chapter he advances two major reasons why the "person in the life space" should be considered as a highly differentiated object. He states that different parts of the body often carry out different activities at the same time and, secondly, that a person cannot be considered a unit, since a change in one need does not change all needs in the same direction and to the same extent. His second reason clearly explains why the "person" of Fig. 1 must be considered as being differentiated. (The differentiation in the life space would, of course, reflect the differentiation in the person.) The first reason — i.e., the possibility that "the person in the life space" is engaging in more than one activity simultaneously — is adequately handled by assuming that the behaving self is in an overlapping situation, a concept much employed by Lewin, or that there is more than one behaving self. Apart from many other considerations, the notion of a differentiated "person in the life space" makes it rather difficult to apply the concept of a force acting upon the person, since the point of application of the force is ambiguous.

Behavior. The term "behavior" has been employed to refer to any change in the life space — i.e., any change which is subject to psychological laws. Thus, not all movements of the person nor all changes in the environment resulting from actions of the person are considered behavior. For example, if a child is moved from an automobile to his home while he is asleep, his movement does not represent behavior; or if a boy hits a baseball and the baseball accidentally breaks a window, the "breaking of the window" cannot be considered to be the boy's behavior. Behavior, in this sense, takes place in the life space rather than in an observable space — or, to use Koffka's (1935) terms, in the behavioral rather than in the geographical environment. It is clear that in this usage "behavior" is not directly observable; it must be inferred. That is, just as attractiveness of any region in the life space must be inferred, so too must change of attractiveness be inferred and, similarly, change of position of the behaving self as well as its initial position must also be inferred. Lewin has also used "behavior" in its more common-sense meaning of the observable interaction between the individual and the objective environment.

The characteristics of the life space (or person) are deduced from observed behavior in an observed environment. However, since we know that many features of the life space tend to remain the same over long periods of time (e.g., a person's enjoyment of tennis on certain types of occasions) despite momentary variations in other characteristics, we can — knowing these characteristics and knowing the observable behavior — make inferences about the objective environment. For example, knowing that

> . . . the people who feel at home in a given community, who seem to fit best, are either highly dependent or authoritarian in personality provides some insight into the characteristics of the community. In a similar way, if, upon investigating the individuals who feel thwarted by a situation, we find that they are the young and ambitious people, the individuals with considerable personal initiative, the nonconformists, we would also be likely to develop some appreciation of the nature of this situation. (Jahoda, Deutsch, and Cook, 1951, p. 46)

The dimensions of the life space. Lewin points out that the life space of the newborn child may be described as a field which has relatively few and only vaguely distinguishable areas. Future events or expectations do not exist; the child is ruled by the situation immediately at hand; he has no conception of his past experiences. The life space, in other words, has no time dimension.

An outstanding characteristic of the change of the life space during development is an increasing differentiation. This includes an increasing differentiation in the *time dimension* (see Fig. 3 for a representation of the time dimension in the life space). Plans extend further into the future, there is an increased tolerance of delay, and activities of increasingly longer duration are organized as one unit. As the individual matures, his picture of the possibilities which exist in the future play an increasingly important role in determining his morale.

The differentiation of the life space brings with it a differentiation of the *reality-irreality* dimension (see Fig. 3). The different degrees of irreality correspond to different degrees of fantasy. They include both the wishes and the fears. A daydream, a vague hope, has in general less reality than an action; an action has more reality than a "talking about it"; a perception is more real than an image; a faraway "ideal goal" is less real than a goal that determines one's immediate action. Dynamically, the level

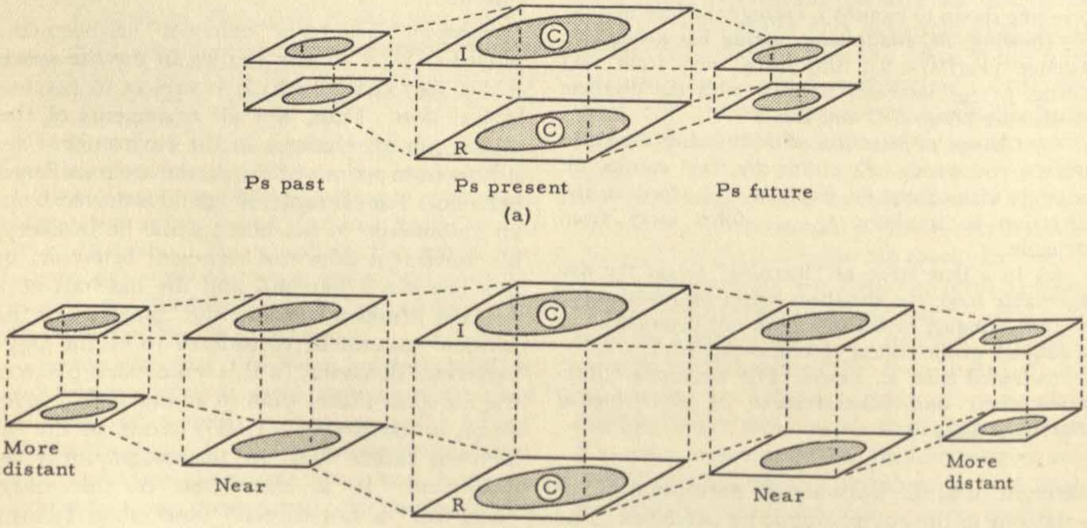


FIG. 3. DIMENSIONS OF THE LIFE SPACE. The life space at two developmental stages. The upper drawing represents the life space of a younger child. The lower diagram represents the higher degree of differentiation of the life space of the older child in regard to the present situation, the reality-irreality dimension, and the time perspective. *C*, child; *R*, level of reality; *I*, level of irreality; *Ps past*, psychological past; *Ps present*, psychological present; *Ps future*, psychological future. (Adapted from K. Lewin, *Field Theory in Social Science*, p. 246.)

of irreality corresponds to a more fluid medium and is more closely related to the central layers of the person. This fact is particularly important for the understanding of the use of play techniques and projective methods in the study of personality.

> . . . The level of irreality in the psychological future corresponds to the wishes or fears for the future; the level of reality to what is expected. . . . Hope corresponds to a sufficient similarity between reality and irreality somewhere in the psychological future; guilt to a certain discrepancy between reality and irreality in the psychological past. In the young child, truth and lying, perception and imagination, are less distinguished than in an older child. This is partly due to the fact that the younger child has not yet developed that degree of differentiation of the life space into levels of reality and irreality which is characteristic of the adult. (Lewin, 1951b, p. 245, or 1946a)

Psychological ecology. Physics treats its space as a closed system of causes and effects in the sense that all physical changes are conceived to be the results of conditions or changes within the same physical space. The matter is not so simple in psychology, since it is clear that the life space is affected by events which originate outside the life space. As a consequence, psychology must take into account physical and social facts which obey nonpsychological laws and which determine the external events that will impinge upon the perceptual motor regions of the person. The problem of taking into account the physical-biosocial world (the nonpsychological milieu) in which the person is immersed and of understanding the way it is transformed into a psychological environment has been termed the task of *psychological ecology* (Wright and Barker, 1950).

The nonpsychological milieu cannot have direct effects upon behavior because its parts are alien to psychological environments, except as its parts get transformed into goals, barriers, and other psychological facts through perception. However, as Wright and Barker (1950, p. 15) point out:

> It is nonetheless true that the nonpsychological milieu affects behavior indirectly. It provides the physical and social raw materials for psychological habitats; and by supplying these resources it goes far to determine just what behavior shall and shall not be possible for all who live within it. It is possible for a [second-grade child in Midwest] to have a psychological habitat that brings about fraternizing with adolescents at school because the Midwest grade and high schools are in the same building. It is [nearly] impossible for [a child in Midwest] to have a psychological habitat that brings about balancing on train rails because Midwest has no railroad.

A characteristic of the nonpsychological milieu is that many of its features which are essential to behavior are relatively constant or predictable through long sequences of behavior. The chair on which you sit, the floor on which your feet rest, the light which enables you to work, remain relatively constant, and hence it is not necessary to adjust your behavior as you work to constantly changing conditions in the nonpsychological milieu.

An important problem for psychology is that of considering what should be included in the nonpsychological milieu and how the included features should be characterized. It seems clear that the interests of the psychologist in characterizing the nonpsychological milieu is not similar to that of other scientists — e.g., the meteorologist, the economist, the physicist, the biologist, the sociologist — concerned with aspects of this milieu. For one thing, the psychologist, as a psychologist, is not interested in the origins or causation of the conditions that go to make up the nonpsychological milieu. He is not concerned with why a given chair will hold only a person weighing less than 200 pounds, but rather with the fact that the chair will not support the "sitting behavior" of a person weighing above this amount. That is, the psychologist is interested in characterizing the various components of the nonpsychological milieu in terms of the perceptions and manipulations they will *support* or permit, the perceptions and manipulations they will *coerce,* and the chain of events in the nonpsychological milieu which will be released by psychological interaction with any of its components.

There are various ways of characterizing the nonpsychological milieu. One can characterize it as Chein (1954, p. 116) has done in terms of "the viewpoint of a psychologically sophisticated observer who presupposes psychological individuals with wants and desires, but who is not at the moment particularly concerned with any one person." Chein distinguishes the following features of the nonpsychological milieu: I. *Stimuli* (anything which is capable of initiating a change in the stream of activity); II. *Goal objects and noxiants* (objects or situations which can serve as satisfiers or which can produce pain or unpleasantness); III. *Supports and constraints* (features of the environment which make particular behaviors feasible or infeasible); IV. *Directors* (features of the environment which tend to induce specific directions of behavior); V. *Global features of the environment* (e.g., its stability, its degree of organization).

The nonpsychological milieu can be characterized in terms of the different kinds of "objects" which compose it. Thus, Parsons and Shils (1951, p. 254) list the following types of objects toward which action can be relevant: the self, others as personalities, collectivities or groups, organisms (including one's own and others'), physical objects, cultural objects (beliefs, symbols, standards, etc.) — and they indicate some of the means of characterizing the action-relevant properties of these different types of objects.

Heider (1953), in a very stimulating unpublished monograph, deals extensively with the question of "changes whose source does not lie in the person but in the environment." Following Angyal's (1941) usage he has termed these changes "heteronomous changes." He characterizes the difference between physical objects and people or groups in terms analogous to the distinction between perception and action. He points out that in perception, the initial focus (the object in the environment) does not guide the causal process through to the final focus (the perception); the fate of the waves emanating from the object is not controlled by and is therefore independent of the object. In action, the initial focus (the plan) guides the causal process to its final focus (the event). Action, conceived of as a certain kind of structure of events in the system comprising the person and the environment, is characterized by an equifinality, the invariance of the end with the variability of the means. The end state is enforced so that a series of independent pushes tends to lead to a unitary effect.

A person (or group) is the cause or the source of heteronomous change in a different sense than is a physical object, since he has greater control over the causal lines emanating from him. Heider indicates that if you meet a person who has certain intentions toward you (suppose he wishes to kill you), it means that your environment contains conditions which are convergently directed toward this state and, if the person has enough power, this state will sooner or later be brought about, whatever you do. Compare this, on the other hand, with avoiding falling stones on a mountain trail; the stones don't change their direction of fall and follow you if you jump out of their way.

In the inorganic world any complex process with specific properties depends on a great many conditions, and it will occur only if all these conditions are met. However, if a person is involved (assuming he has sufficient power) there is only one condition necessary for an event to occur: an appropriate intention. The fact of personal causality, that a person can bring about an effect if he intends to do so regardless of widely varying circumstances, makes people uniquely significant components of the nonpsychological milieu. Physical objects do not "wish" to nor are they able to create events which have as their terminal focus a specified change in an individual's life space. Only people can have the desire and power to determine the contents of an individual's psychological environment. Since the human being is from birth the terminal focus of actions of others, it is evident that the understanding of his behavior requires that he be understood in his interrelations with others.

Wright and Barker (1950) have described a useful concept for characterizing those parts of the nonpsychological milieu which tend to "coerce" similar psychologial environments or habitats in different people. They use the term "behavior setting" to refer to any part of the nonpsychological milieu that is generally perceived by the people of a community as appropriate for particular kinds of behavior. Some examples of a behavior setting are: a doctor's office, a social party, a classroom, a free play period, a fishing trip, a movie, the status position of a schoolteacher, breakfast, dressing, the status position of a father, homework, Fourth of July celebration. Behavior settings include positions in a social structure (e.g., boss, foreman, and worker) as well as regions which are more easily visualized in terms of geographical locus.

It is evident that behavior settings are "coercive" primarily because there exist shared frames of reference or collective intentions as to what is appropriate behavior in a given behavior setting or with a given behavior-object. As defined by Wright and Barker, a behavior setting always has two sides. On the one side, there is always a part of the nonpsychological milieu, which may be physical, social, or both, as in the respective instances of a screwdriver or a vacant lot, a traffic law or a social clique, and a writ of *habeas corpus* or a school classroom. On the other side, there is always a social norm or a shared frame of reference that defines the setting as one which is appropriate for particular kinds of behavior.

It would be misleading if our discussion in the foregoing paragraphs led to the interpretation that a nonpsychological milieu would coerce different people to have similar psychological reactions only if it has interpersonal significance. The fundamental identity in the functioning of our respective nervous systems "coerces" individuals with similar experiences who are exposed to similar environments to perceive and to react similarly. As Asch (1952, p. 129) has pointed out, we are able to enter relations with others only because

> . . . the space that surrounds us possesses the same directions for all people; the structure of time — its extension into the past and the future, the order of events that mark off its periods — is, or can be, the possession of all; the relations of cause and effect are similarly perceived. When we act with others we do so on the basis of this objective, open field that surrounds us and which stands in a similar relation to each of us.

STRUCTURAL CONCEPTS

Topology may be defined as a branch of geometry which investigates the properties of figures that remain unchanged under continuous transformation or "stretching." These are the qualitative relationships of connection and position. Early in his career, Lewin felt that topology might be particularly fitted to the special problems of psychology. The concern of Gestalt theorists with part-whole relationships, with "belongingness" and "membership character" suggested the relevance of topological concepts to psychological questions.

Although topology is a young mathematical discipline which has been developed almost exclusively within the past 50 years, it is now a complex and high-powered mathematical tool. The "topological concepts" employed by Lewin are, from the point of view of mathematical topology, of a rudimentary nature and are torn out of their mathematical context. Mathematicians have sometimes seen little relationship between the "topological concepts" employed by Lewin and mathematical topology. Lewin, however, has written in implicit answer to mathematical objectors that:

> The mathematician will have to realize . . . that to apply a system of mathematical concepts in an empirical field one does not necessarily have to prove directly the adequacy of the basic mathematical axioms of this system one by one. It is well to prove the fruitfulness of some of the desired propositions of this mathematical system for the representation of the empirical properties of the field in question. If the representation of spatial relations in physics by Euclidean geometry had not been permitted until its axioms (such as the divisibility *ad infinitum* of any part of the space) were proved one by one to hold also for the physical space, physics could never have used Euclidean geometry. All one can say is this: if one coordinates certain physical processes to certain geometrical entities one can make certain physical predictions . . . Exactly the same procedure is followed if certain psychological processes (such as social locomotion) are coordinated to certain entities of topological or hodological geometry (such as path). There can be no other meaning and no other proof of the applicability of these geometries to psychology than the fruitfulness of predictions based on such coordination. (Lewin, 1951b, p. 22, or 1940)

Lewin's answer does not adequately meet the objections of his mathematical critics; they feel that Lewin, although employing some of the terms of topology, has not used any of the theorems of mathematical topology. However, regardless of the mathematical rigor of Lewin's "topological" concepts, they serve an important function in his theoretical system. In general, one may say that their function is to allow one to determine which events are possible in a given life space and which are not. Dynamic concepts are necessary, in addition to those of topology, in order to determine which of the possible events will actually occur in a given case.

Let us illustrate Lewin's use of topological concepts to characterize "which events are possible" by reference to Fig. 4, which depicts the life space of a student who is considering how to spend his evening. Let us assume that he envisages only four possibilities: seeing a movie, visiting a friend in the hospital, writing a term

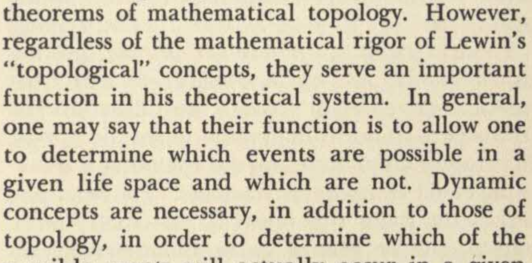


FIG. 4. TOPOLOGICAL REPRESENTATION OF THE PSYCHOLOGICAL SITUATION OF A STUDENT CONSIDERING FOUR WAYS OF SPENDING HIS EVENING. The individual, *P*, is in the position of choosing among four activities: visiting a sick friend, a_1; writing a paper, b_1; seeing a movie, c_1; or earning some money, *d*. He sees the path to a_1 as involving calling the hospital, a_3, and traveling to the hospital, a_2. The path to b_1 is twofold, consisting of either going to the library, b_4; reading at the library, b_3; taking notes at the library, b_2; or of working with the books and notes already at home, b'. The path to the goal of seeing a movie is seen to require a trip downtown, c_3, and the buying of a ticket, c_2. The individual perceives no path leading from his present position to region *d*.

paper, and earning some money; that is, only these alternatives exist psychologically and, therefore, only they are part of the life space. From the diagram, one can further see that for this particular student "going to a movie" requires that he "go downtown" and "buy a ticket to the movie" before he "sees the movie." "Seeing a movie," in other words, is only possible if it is preceded by the activities of "going downtown" and "buying a ticket." We would expect that if by chance he found that he did not have the money to buy the ticket after going downtown that he would not see the movie, since the only region connected with "seeing the movie" is "buying a ticket." The regions of "sneaking into the movie" or of "getting someone else to buy his ticket" do not exist psychologically and, thus, if he cannot buy a ticket, it is impossible for him to see the movie. The other alternative ways of spending the evening could be similarly discussed. For example, Fig. 4 indicates that even if the library were closed, another way of going about writing his term paper exists as a possibility. On the other hand, for this student, it does not make sense to travel to the hospital to visit his friend without first calling up to see whether or not visiting is permitted. It is also evident that the activity of "earning money" will not be engaged in, since there doesn't exist, psychologically, a linkage or path between what the student is now doing and this particular activity. In other words, he doesn't see how he can get to "earning money" from his present position.

Let us employ Fig. 4 to illustrate some of Lewin's specific topological concepts. One of his basic concepts is that of *region,* which may be defined as any distinguishable part of the life space (or person). Regions of the psychological environment refer to present or contemplated activities rather than to the objective areas in which activities occur or to the external objects with which activities are linked (Leeper, 1943, pp. 92-95). A child who is using the back porch as a stage and a chair as a throne is in the region of "play acting" (on the porch with a chair as a throne), rather than in the region of the "porch." It should be obvious, however, that the characterization of any activity will be facilitated by knowledge of its time and locale and of the objects with which it is concerned. In Fig. 4, all of the areas enclosed by a line represent regions (e.g., the region of considering how to spend the evening, the region of going downtown, the region of buying a ticket). Some regions include subregions (e.g., visiting one's friend in the hospital may include subregions of giving him a persent, encouraging him, amusing him). However, psychological activities are not thought to be infinitely divisible. The degree of *differentiation* of a region refers to the number of subparts within a region. Any region which has no distinguished subparts may be called a *cell.*

The *position* of the behaving self, *P,* is defined by the activity region in which it is located. Any change of position of any region within the life space is conceived to be a *locomotion.* The term "locomotion," however, is primarily used in reference to locomotion of the behaving self rather than in reference to locomotion of parts of the psychological environment. In Fig. 4, a locomotion would take place if *P* moves out of the region of "considering how to spend his evening" into any other region, or if some rearrangement of the psychological environment took place so that, for example, the student now views "visiting his friend" as leading to, rather than away from, "seeing a movie." (This might occur if he called up the hospital and found out that they were showing movies to the patients and their visitors.)

Locomotion from one region to another involves movement of the behaving self from its present to its terminal position through a path of *neighboring regions.* Two regions are neighboring if both regions have a common boundary and are otherwise foreign to each other. The *boundary of a region* consists of those cells in the region for which there is no surrounding boundary that lies entirely within the region. Regions $A_1, A_2, \ldots, A_n$ are said to form a *path* if A_1 is touching A_2, A_2 is touching $A_3, \ldots$, A_{n-1} is touching A_n. In Fig. 4, there is no psychologically existing path between "considering how to spend the evening" and "earning some money," while there are two possible paths between "considering how to spend the evening" and "writing a paper." The *distinguished path* between any two regions is the path along which the individual anticipates that he will locomote if he chooses to proceed from one region to another. It is, in other words, the preferred or psychologically best path.

As we have pointed out in an earlier section, Lewin developed a geometry he termed "hodological space" to represent direction in the life space. As in other geometries, the concept of direction is based upon the concept of distinguished path. However, it is evident from the preceding paragraph that the concept of distinguished path in hodological space takes into consideration the cognitive structure and dynamic properties of the life space. A change in the psychological environment (i.e., in the cognitive structure) or of the attractiveness of its various features is a sufficient condition for changing the direction between regions.

Direction at *A* toward *B* is defined as the relation between the regions *A* and *B* which is determined by the beginning step of the distinguished path from the region *A* to *B*. Four types of direction of locomotion can be distinguished (the definitions which follow represent an adaptation of Leeper's modifications of Lewin's original definitions): (1) *direction toward* — any locomotion from one region to another which increases the force toward or decreases the distance from the terminal focus of locomotion (or goal-region); (2) *direction away* — any locomotion from one region to another which decreases the force toward or which increases the distance from the terminal focus of locomotion; (3) *opposite direction* — two paths have opposite direction if the distinguished path for locomotion toward one goal-region corresponds with the most unfavorable path for locomotion toward the other goal-region; and (4) *direction at "right angles"* (Leeper, 1943, p. 205) — a path is at right angles to another path if locomotion along it does not change the strength of force toward the terminal focus of the other path or if locomotion along it does not change the distance to the terminal focus of the other path.

In addition to the preceding types of direction, Lewin distinguishes between *equal* and *partly equal* directions. If one compares two paths from one region to different goals which share the same first step (s), the degree of equality of direction at the starting point approaches zero if these paths show a high degree of psychological unity (all steps in the path tend to be fused as part of an inseparable whole) as well as if each shows a very low degree of unity (so that each step is viewed as relatively separate and independent from the other steps in the path). Forces with partly equal directions summate, but not as completely as forces in fully equal directions.

Without elaborating, let us point out that Lewin's discussion of the psychological unity of a path requires that a path, such as the chain of regions *A-B-C-D-E* connecting *A* and *E*, be conceived as a series of *overlapping paths* which consist of *A-B, A-B-C, A-B-C-D,* and *A-B-C-D-E.* If the *potency* of *A-B-C-D-E* is high then there may be no distinguishable subunits on the path. If, on the other hand, the potency of any subpath is high, the total path *A-B-C-D-E* may not exist as a psychological unit (for a discussion of the concepts of *overlapping situations* and *potency* see several paragraphs below).

While the factor of *psychological distance* is of crucial importance in studying problems of the strength of forces, Lewin developed the concept of distance only sufficiently to indicate that one path is longer than another if it includes the other path within itself; however, the fact that one path is longer than another does not imply that it includes the shorter path within itself. However, in a subsequent *Analysis of the Concepts: Whole, Differentiation, and Unity* (Lewin, 1951b, or 1941), he defines the minimum distance between any two cells as the minimum number of boundaries crossed by a path from one of these cells to the other. This later definition, which has also been employed by Bavelas (1948) in a modified form, runs into the difficult question of whether it is fruitful to consider all cells (or undifferentiated psychological activities) as being of the same size.

We have indicated in a preceding paragraph that direction in the life space is dependent upon the *cognitive structure.* Thus, if the individual has no clear knowledge of the sequence of steps necessary to achieve a given objective (e.g., to solve a mathematical problem, to influence a policeman not to give him a speeding ticket, to escape from a burning building), the direction of locomotion necessary to obtain his goal is not known. Most new situations are cognitively unstructured, since the individual is likely to have no knowledge of "what leads to what." Hence it follows that in such situations the individual is unlikely to know what direction he should follow to achieve his objective. As he strikes out in any given direction he does not know whether it is leading him toward or away from his goal. His behavior will be exploratory, trial-and-error, vacillating, and contradictory rather than parsimonious. If reaching the goal has positive significance and not reaching it has negative significance for the individual concerned, then being in a region that has no clear cognitive structure results in psychological conflict, since the direction of the forces acting upon him are likely to be both toward and away from any given region. There will be evidences of emotionality as well as cautiousness in such situations. In addition, the very nature of an unstructured situation is that it is unstable; perception of the situation shifts rapidly and is readily influenced by minor cues and by suggestion from others.

This characterization of some of the psychological implications of cognitively unstructured situations is of considerable usefulness to social psychology. For example, Lewin has employed this concept to give insight into the situation of the adolescent (1939a or 1951b). He pointed out that the change from child to adult is a shift to a more or less unknown position:

> . . . Entering a new social group can mean something very similar to being thrown into a cognitively unstructured field, being forced to stand on unfirm ground and not knowing whether the "right thing" is being done. The uncertain character of the adolescent's behavior and his conflicts can partly be explained by the lack of cognitive clarity concerning the adult's world which he is going to enter. It follows that this uncertainty is greater the more the individual has previously been kept out of the adult world and has been kept in the dark about it . . . [In addition] the time of sexual maturity brings with it changes which make this part of the life space, which is so close and vital to the individual, strange and unknown. In this case the change does not mean merely the usual uncertainties of a new and strange environment; but, in addition, a region which previously appeared to be well known and reliable becomes now unknown and unreliable. This change necessarily shakes the belief of the individual in the stability of the ground on which he stands and perhaps even in the stability of the world at large. Since the region of the body happens to be very important and central for everyone, this doubting might be rather fundamental. (Lewin, 1951b, pp. 138-139, or 1939a)

The concept of lack of cognitive structure has also been fruitfully applied to characterize the situation of members of an autocratic group where decisions are arbitrarily made (or made on a basis which is unclear to the members) by the person in authority. It also has been used to give insight into the psychological situation of minority group members (Lewin, 1948a, or 1935a), of people suffering from physical handicaps (Barker, Wright, Gonick, 1946), of the *nouveaux riches*, or of other persons crossing the margin of social classes. It may be applied to any situation in which the consequences of behavior are seemingly unpredictable or uncontrollable, in which benefits and harms occur in an apparently inconsistent, fortuitous, or arbitrary manner, in which one is uncertain about the potential reactions of others to being in a certain activity region.

The concept of *overlapping situations* has been employed to characterize instances where the person is on more than one path at the same time. The simplest example is that of divided attention: a student eating his lunch and trying at the same time to study for an examination. One could analyze most instances of experienced conflict as being examples of overlapping situations — e.g., the role conflict experienced by some career women between being "professional" or being "feminine" at any particular instant.

Barker, Wright, and Gonick (1946) have pointed out five aspects of overlapping situations: (1) *Consonance*. The extent to which the overlapping situations require activity that is more or less congruent is called their consonance. Overlapping situations can vary from situations leading to identical behaviors (i.e., paths have same directions) to situations requiring completely antagonistic behavior (i.e., paths have opposite directions). (2) *Potency*. The influence of one situation relative to all simultaneously acting situations is called the relative potency of the situation. A person in the overlapping situations of "reading the paper" and "conversing with wife," in which the latter situation is of relatively higher potency, will predominantly converse with his wife, with perhaps an occasional glance at the headlines or a slight feeling of annoyance. (3) *Valence*. An important characteristic of overlapping situations is their relative valence or desirability. (4) *Barriers*. A fourth characteristic of overlapping situations is the nature of the perceived barriers or obstacles in each of the situations. (5) *Extent of common parts*. This characteristic refers to the extent of the part common to the overlapping regions.

While the representation which is employed in topological diagrams such as Fig. 4 is useful in depicting the possibilities which exist in the life space, it obviously does not tell us which of the alternative possibilities will be chosen. The selection among the possibilities will be determined by such dynamic factors as the relative strength of the goals of "seeing a movie," "visiting the friend in the hospital," "writing the paper"; the attractiveness of the activities which are necessary to reach each of the different goals; the probabilities that the goals will be reached by the available paths.

Lewin's "topological" representation (of which Fig. 4 is a typical example) is most adequate for describing the activities that an individual perceives as being necessary for the attainment of a given objective and, thus, for depicting alternative courses of action as they appear in the life space of the acting individual at the planning or decision-making stage. It should be made clear that the relations represented in a topological diagram can be translated without loss of exactness into statements. However, the diagram is often a more convenient way of expressing the relations among a limited number of facts than would be a series of statements. On the other hand, the topological diagram is very awkward and, at times, even a hindrance when one attempts to depict certain types of perceived possibilities. This is true when the means-ends relations are numerous or complex. Moreover, Heider (unpublished manuscript) and Leeper (1943) have pointed out that the "topological" diagram is inconvenient when one wishes to represent the *sufficient* rather than the *necessary* conditions for an event. For example, a child may perceive "eating jam" as being a sufficient but not a necessary condition for being punished. The fact that "eating jam" irrevocably leads to punishment in the child's view is difficult to represent without the use of arbitrary devices; on the other hand, there is no difficulty in representing high school graduation as a necessary condition for entering college. Heider has also pointed out that the "topological" diagram is inconvenient when the perceived source of change in the life space is not the actor but is somebody else or an impersonal source. Thus, it is difficult to represent topologically the expectation that "if my boss gives me a raise, I will be able to buy a new car." Similarly, it is difficult to represent topologically instances where one is the recipient of an event — e.g., that a child expects to have a good time when his grandparents visit.

In addition to the foregoing difficulties often encountered in "topological" representation, Lewin's attempt to go beyond topology to treat problems of distance and direction in a manner suitable for psychology in his "hodological space" was not mathematically successful. For example, although the concept of distance is basic to the definitions of direction which are employed in hodological space, it was never more than partially defined. Lewin was well aware of the provisional nature of "hodological space" and in his last years was engaged, with the aid of some of his students and some mathematicians, in attempts to develop a more adequate geometry. Bavelas (1948), one of the most ingenious of Lewin's students, developed a mathematical model which has considerably more mathematical rigor to it than "hodological space." Instead of defining the distinguished path in terms of dynamic factors, the criterion of a distinguished path is purely a positional one — that of "minimal regions to be crossed by a path from *a* to *b*." Employing five basic assumptions and a series of definitions developing from these assumptions, Bavelas is able to give unequivocal meaning to a number of concepts which seem to have applicability to personality and group theory: "structures," "outermost region of a structure," "innermost region of a structure," "central region of a structure," "peripheral region of a structure," "maximum" and "minimum" distances within a structure, "steps toward or away, or neither toward nor away from a specified cell," "a straight line path," etc. Some of these concepts have already been employed fruitfully in experimental studies of different types of communication structures in groups (Leavitt, 1951). Harary and Norman (1952), working at the Research Center for Group Dynamics, have employed the mathematical theory of graphs (a branch of combinatorial topology) to the study of communication networks and group structures.

Let us conclude this section by indicating that although Lewin's "topological" and "hodological" concepts are not much more than diagrammatic representations of his brilliant theoretical and experimental insights, it would be foolish to neglect the suggestive value of his imaginative attempt to develop a geometry suitable for psychology. He pointed out the need for a new mathematics based on axioms different from those of the mathematics developed for the physical sciences. He indicated some of the properties that would be required of a geometry adequate to handle psychological space and he stimulated a more widespread interest in the development of such a geometry.

DYNAMIC CONCEPTS

In Lewin's system, dynamic concepts have the function of enabling one to determine which of the possible psychological events will occur. Just as his geometrical concepts are not independent of dynamic factors, so too his dynamic concepts are defined in a way which presupposes an adequate geometry. However, despite this interdependence of concepts in his formal "topological and vector psychology," many of Lewin's basic dynamic concepts were formulated prior to his explicit use of topological concepts.

The concept of tension system. The beginnings of Lewin's dynamic conceptions can be traced back to his controversy with Ach. Ach (1910) had conducted a series of experiments concerned with the learning of paired nonsense syllables in which the learned association between the previously paired nonsense syllables (e.g., between rik-tik) was put in conflict with the test instructions which called for a different type of response than that required in the learning situation. For example, the subject would be asked to invert the stimulus syllable and thus be required to respond to "rik" with "kir" instead of the "tik" which he had learned to associate with "rik" during the initial learning situation. Ach found that reaction time in the test situation was influenced by whether the "associative tendency" acquired in the initial learning coincided or conflicted with the "determining tendency" resulting from the instructions in the test situation. While Ach's experiments convincingly demonstrated conflict and delayed reaction time, Lewin argued that the conflict was not between "association" and a determining tendency but between two determining tendencies. It was his belief that there is no "force" within mere association to lead to reproduction, that reproduction itself must be motivated.

To prove his contention, Lewin conducted a series of experiments (Lewin 1917, 1922), that were much like Ach's except for a few minor modifications. They demonstrated that the existence of an association is not, *per se,* sufficient for recall of the associated syllable on presentation of the stimulus syllable and, thus, for interfering with a different response to the stimulus syllable. He reaffirmed that an association (or, to use his terms, a cognitive structure) does not provide the "motor" for mental activity. A tension system is always necessary for activity, including the activity of reproducing previously learned nonsense syllables.

The concept of *tension system* plays a central role in Lewin's theorizing, and it has been the stimulus to a series of ingenious experiments on recall and resumption of interrupted activities, substitute activities, and satiation. It is most appropriately applied in connection with the person as represented in Fig. 1 (see p. 190). A system in a state of tension is said to exist within the individual whenever a psychological need or an intention (sometimes referred to as a quasi-need) exists. Tension is released when the need or intention is fulfilled. Tension has the following conceptual properties: (a) it is a state of a region or system S which tries to change itself in such a way that it becomes equal to the state of its surrounding regions $S^1, S^2 \ldots, S^6$, and (b) it involves forces at the boundary of the region S in tension.

There exists a definite relation between tension systems of the person and certain properties of the psychological environment. In particular, a tension may be related to a positive valence for activity regions in the psychological environment which are perceived as tension-reducing, and a negative valence for the region in which the behaving self is at present. However, the existence of a region of positive valence (a goal region) depends not only upon the existence of tension but also upon whether or not there are perceived possibilities for reducing the tension. The individual who has no conception of how to reduce a tension, and who therefore perceives no goal region, will have his behavior determined by his desire to leave his present negative region for another, and to leave that for another, and so on. His behavior will be characterized by locomotion in the direction away from his present region — or, in observable terms, by restless movement. Search behavior, occurring in the situation where the individual knows a goal region exists but does not know its location, has many similarities to restless activity but is distinguished from it by the fact that the direction toward the goal is dominant and the leaving of the present region is only a means to that end (Lewin, 1938, pp. 62-63).

When a goal region which is relevant to a system in tension exists in the psychological environment, one can assert that there is a force acting upon the behaving self to locomote toward the goal. A tension for which there is a cognized goal leads not only to a tendency to actual locomotion towards the goal region but also to thinking about this type of activity. This may be expressed by saying that the force on the person toward the goal exists not only on the level of doing (reality) but also on the level of thinking (irreality).

From the foregoing assumptions about systems in tension it is possible to make a number of derivations. Thus, it follows that the tendency to recall interrupted activities should be greater than the tendency to recall finished ones. Zeigarnik (1927) and many others have conducted experiments in which subjects are given a series of tasks to perform and are then prevented from completing half of them (selected at random). Later, the subjects are asked to recall what tasks they had performed. The results are presented in the form of a quotient, commonly called the Zeigarnik quotient (ZQ):

$$\frac{\text{unfinished tasks recalled (RU)}}{\text{completed tasks recalled (RC)}}.$$

Zeigarnik predicted a quotient of greater than 1. The obtained quotient was approximately 1.9, clearly supporting Lewin's assumptions. However, since many completed tasks were also recalled, it was obvious that additional factors were involved. Analyzing the situation of the subject at the moment of recall, she concluded that in addition to the force on the person to think about, and hence to recall, the incompleted tasks, there was also present a force to recall *both* incompleted and completed tasks *induced* upon the person by the experimenter's instructions to "try to recall the tasks you worked on earlier." The Zeigarnik quotient could be viewed as a function of the relative strengths of the induced force to recall all tasks and of the force to recall the incompleted tasks. As the strength of the induced force increases in relation to the force toward the task goals, the quotient should approach 1; as it decreases in relative strength, the quotient should increase beyond 1. These additional predictions, which follow from an analysis of the situation at recall, were borne out in experiments by Zeigarnik and others. Thus, if the strength of motivation associated with the interrupted task is relatively high, or the strength of the experimenter's pressure to recall is low, or if the task is interrupted near its end, the Zeigarnik quotient will be high.

A number of more recent experiments have indicated that the situation of recall is frequently even more complex than indicated above. When not finishing a task is interpretable as a personal failure (e.g., in an experiment where the tasks are presented as measures of a socially esteemed ability) and when recall of failure threatens one's self-esteem, or when the recall of success raises a lowered self-esteem), the Zeigarnik ratio tends to be less than one. For a more thorough discussion of the effects of self-esteem involvement on the Zeigarnik quotient see Alper (1946, 1948), Glixman (1948), Prentice (1944), Rosenzweig (1943), and Zeller (1950).

It should be noted that the terms "unfinished" and "completed" in the Zeigarnik quotient should refer to the subject's view of the tasks rather than to whether or not they were objectively completed. If the subject achieves his goal in relation to the task even though objectively it is unfinished, one would expect no force toward recall and, contrariwise, if the subject's goal is not attained even though the task is completed, one would expect the tension to persist. Data from Zeigarnik's experiment and from an experiment by Marrow (1938) support the notion that "psychological" reality rather than "objective" reality is what is significant here in predicting an individual's behavior.

From the general assumption that when there is a need present for a certain goal there corresponds a force causing a tendency to locomote toward the goal, it follows that as long as the tasks are psychologically incomplete the subject should continue to try to perform the task (providing there are no sufficiently strong counter forces). Ovsiankina (1928) created conditions such that the subject thought he was interrupted by chance and at a later time he was left free to do as he wished in an experimental room. There was 100 percent resumption of the interrupted task. When the interruption appeared to be intentional, there was still 82 percent resumption.

Since the construct of tension refers to the state of one system relative to the state of surrounding systems, it presupposes a geometric representation of the person and a distinction between functional subparts or "systems" within the person. Lewin (1935b), in conceiving of the person as a series of differentiated subsystems, is rejecting the catch phrase "everything is related to everything else" as misleading, since there is a great range of interdependence between the subsystems of the person. Much of the research of Lewin and his students is concerned with the question of what factors determine the interdependence of tension systems and with the "fate" of tension with varying conditions of action.

Lewin postulated that the change in the tension difference between any two tension systems would depend upon the time interval and the degree of interdependence of the two systems. With increased time, one would expect the tension differences between interdependent systems to decrease. Thus, Zeigarnik in her experiments on the recall of interrupted activities and Ovsiankina in her experiments on the resumption of interrupted activities both found that as time elapses there is a decreasing tendency to recall or resume the interrupted activities.

The degree of interdependence of two tension systems is conceived to be a function of the degree of *fluidity* or rigidity of the person and of the structural relations or connections between the two systems. The fluidity of the person is affected by general states of the person. For example, a person is considered to be more fluid when tired, when he is undergoing a wave of emotional tension, when he is young, when he is in a "make-believe" or irreal situation, when he is more intelligent, etc. Experiments by Zeigarnik have indicated that in more fluid states, the tension differences between systems tend to dissipate more rapidly.

Experimental studies of rigidity by Kounin (1941), in which he compared three groups of subjects of equal mental age (an old feebleminded group, a young feeble-minded group, and a normal group), led him to conclude that

> . . . any performance which requires a certain degree of communication between neighboring regions—, is to such an extent made difficult for the older and/or more feeble-minded individual. The task may be predominantly of a cognitive

> nature . . ., or a motor nature, . . . or of a volitional nature. If a task is facilitated by a lack of communication between neighboring regions, such a task will be more efficiently and accurately performed by an older and/or more feeble-minded individual.

The structural relations between systems refer to such questions as whether or not two systems are subparts of a larger system, whether or not two systems have a relationship of simple dependence, interdependence, or organizational dependence, and the proximity of or distance between regions. While the problem of characterizing structural arrangements in logically strict conceptual terms has received only scant attention (Bavelas, 1948), (Lewin, 1951b, Appendix), a number of interesting experiments have been conducted which deal with structural relations. The experimental studies of *substitution* are relevant in this connection.

Lissner (1933) initiated the study of the value which one activity has for reducing a tension originally connected with another activity by a technique of resumption; some of the other experimenters have employed the technique of recall. The substitute value is measured by the amount of decrease in resumption or recall of the interrupted original activity after a substitute activity has been completed. The results of the experiments on substitute value can be summarized briefly as follows: (1) Substitute value increases with the perceived degree of similarity between the original and the substitute activity and with the degree of difficulty of the substitute activity (Lissner, 1933). (2) Substitute value increases with increasing temporal contiguity between the original and the substitute activity and with the attractiveness of the substitute activity (Henle, 1942). (3) The substitute value of an activity (e.g., thinking, talking, or doing) depends upon the nature of the goal of the original task. Tasks that are connected with the goal of demonstrating something to another person (e.g., the experimenter) require an observable substitute activity (not merely "thinking" without social communication); "realization tasks," in which the building of a material object is the goal, require "doing" not only telling how it can be done; for intellectual problems, talking (or telling how it can be done) can have a very high substitute value (Mahler, 1933). (4) "Magic solutions," "make-believe solutions," or solutions which observably violate the requirements of the task have little substitute value for tasks at the reality level. However, if the situation is a make-believe or play situation, make-believe substitutions will have substitute value (Sliosberg, 1934; Dembo, 1931). (5) A substitute activity which is identical with the original activity will have little substitute value if it does not serve the same goal. Thus, building a clay house *for* John will have little substitute value for building a clay house for Mary. If the emphasis is on building a clay house and not upon the *for somebody,* then, of course, substitution will occur (Adler and Kounin, 1939). (6) Having somebody complete a task that the subject is interrupted in tends to have little substitute value, particularly when completion of the task by oneself is related to self-esteem. However, when pairs of individuals work cooperatively on a task, the completion of the task by one's partner has considerable substitute value (Lewis, 1944, and Lewis and Franklin, 1944).

The research findings with respect to substitute value have implications for a wide range of problems in psychology — from the relative gratification value of individual versus socially shared projective or fantasy systems to the development of specialized roles within a group. Let us briefly illustrate with the very important finding that the actions of another person can be a substitute for one's own actions if there is a cooperative relationship. The fact of substitutability enables individuals who are working cooperatively on a common task to subdivide the task and to perform specialized activities, since none of the individuals in a cooperative situation has the need to perform all the activities by himself. In contrast, the individual in a competitive situation is less likely to view the actions of others as substitutable for similarly intended actions of his own. Thus, when a competitive situation exists in a group, there is less likely to be a development of specialization in activities (Deutsch, 1949a, b).

The concept of tension systems has also been fruitfully employed in experimental studies of *satiation.* One can distinguish in regard to most needs a state of hunger, of satiation, and of oversatiation. These states corresponds to a positive, a neutral, and a negative valence of the activity regions which are related to a particular need or tension system. Karsten (1928) has studied the effect of repeating over and over again such activities as reading a poem, writing letters, drawing, and turning a wheel. The main symptoms of oversatiation appear to be (1) the appearance of subunits in the activity leading to the disintegration of the total activity and a loss of meaning of the activity; (2) increasingly poorer quality and greater frequency of errors in performing the task; (3) an increasing tendency to vary the nature of the task, accompanied by a tendency for each variation to be quickly satiated; (4) a tendency to make the satiated activity a peripheral activity with an attempt to concentrate on something else while doing the task — this attempt is usually not completely successful and the mind wanders; (5) increasing dislike of the activity and of similar activities with an increased valence for different tasks; (6) emotional outbursts; (7) development of "fatigue" and similar bodily symptoms which are quickly overcome when the individual is shifted to another activity.

Satiation occurs only if the activity has, psychologically, the character of marking time or of getting nowhere. If the activity can be viewed as making progress toward a goal, the usual symptoms of satiation will not appear. Embedding an activity in a different psychological whole so that its meaning is changed has practically the same effect on satiation as shifting to a different activity. The rapidity with which satiation occurs depends upon (1) the nature of the activity — with increasing size of its units of action and with increasing complexity satiation occurs more slowly, (2) the degree of centrality of the activity — other things being equal, activities which are of more significance to the person are more quickly satiated than peripheral activities, (3) the state of the person — the more fluid the state of the person the more quickly he is satiated. The rate of satiation and cosatiation of similar activities (i.e., the spread of satiation effects from one activity to similar activities) decreases with age and with lack of intelligence (Kounin, 1941).

Most of the phenomena of satiation can be explained by assuming that continued performance of a task leads to a lowering of the tension level in the system corresponding to the task; with increasing repetition the system corresponding to the task may be represented as reaching a *lower* level of tension than the surrounding systems. This should lead the person to turn away from the repeated task to other activities. With time, the person will also turn away from activities which are related to tension systems that are interdependent with the satiated system and turn toward activities which are not so related to the satiated activity. Thus, the tendency is to turn away from similar activities and to turn toward dissimilar activities. The breakdown of the satiated activity into disorganized units fits in very well with Lewin's analysis of the "differentiation and unity of a whole based on simple dependence" (Lewin, 1951b, pp. 305-335, or 1941).

The concept of tension system has also been employed to give insight into some of the effects of *frustration.* Barker, Dembo, and Lewin (1941) in their study of *Frustration and Regression* have indicated that psychological development can be characterized as resulting in the following structural changes: (a) an increased differentiation of the person into subsystems, which is reflected in a wider variety of behavior, emotions, skills, knowledge, etc.; (b) a change of organization of the person so that subsystems of the person are hierarchically related to one another rather than related to one another by simple interdependence. Thus, as the individual grows older some systems become more dominant and other systems serve as their "tools"; in infancy, interdependence of systems is more characterized by the simple diffusion type of influence. Barker, Dembo, and Lewin believed that the structural changes brought about by psychological development could be reversed, at least momentarily, by creating a high level of tension in a subsystem of the person and by not permitting this tension to be reduced through activity toward a relevant goal. Presenting a frustrating situation, i.e., one in which an individual is prevented from reaching a desired goal, is one way of creating such tension.

To study the effects of frustration, they created an experimental situation where children (ages 3 to 5) who had been allowed to play with some extremely desirable toys were taken to another part of the room which separated them from these toys by a wire partition. Here they were allowed to play with less desirable toys. Their behaviors, including the constructiveness of their play activities, were observed and compared during the period when they could play with the desirable toys (free play period), during the period when they could not play with the toys (frustration period), and during a later period when they were again allowed to play with the desirable toys (the post-frustration period). The results of the experiment indicate that frustration produces a change in the constructiveness of play activities (i.e., of the degree of organization and differentiation in the play activities) equivalent to an average regression of 17.3 months of mental age.

The dedifferentiation which was produced by the experiment is explained in terms of two related notions, which are developed at some length by Lewin (1951, Appendix, or 1941): (1) The degree of differentiation of a whole (e.g., the person) is inversely related to the strength of tension when this passes beyond certain limits that are determined by the strengths of boundaries of the subsystems comprising the whole, and (2) a decrease in the variety of behavior must also occur if a part of the whole is kept in a fixed state. The amount of decrease depends upon the extent of the fixed areas, their degree of centrality, and their divergence from the normal level. Frustration keeps a certain portion of the person in a state of more or less permanent tension and the variety of behavior therefore decreases. The change in degree of organization of play activities results from a tendency for diffuse spreading of tension when tension is high and thus a bypassing or disruption of organizational relations and, in addition, from the conflict for control over the perceptual-motor regions between the "frustrated" tension system and the tension system related to playing with the less desirable toys.

Let us conclude our discussion of tension systems by pointing out that the concept of tension system and the various experimental studies on the recall and resumption of interrupted activities, substitution, satiation, and frustration have direct relevance to many problems of social psychology (although, as yet, their relevance has not been fully exploited). The concept of tension system is applicable to socially derived needs and intentions, to motives which develop from belonging to a group and from participating in group activities, to interpersonal influences, etc. The various effects of tension upon psychological processes provide criteria for developing insights into the social and group factors which produce individual motivation as well as insight into the social and group factors which facilitate the reduction of individual tensions.

Dynamic properties of the psychological environment. The construct *force* characterizes, for a given point of the life space, the direction and strength of the tendency to change. The combination of a number of forces acting at the same point at a given time is called the *resultant* force. Whenever a resultant force (different from zero) exists, there is either a locomotion of the behaving self in the direction of that force or a change in cognitive structure equivalent to this locomotion.

Lewin distinguishes between *driving* and *restraining* forces. Driving forces correspond to a relation between at least two regions of the life space: the region of present activity and the region of a goal. Driving forces tend to lead to locomotion, hence the statements in the preceding paragraph, strictly speaking, refer only to driving forces. Restraining forces, as such, do not lead to locomotion, but they do influence the effect of driving forces. Any region which offers resistance to locomotion, i.e., any *barrier* to locomotion, is characterized by restraining forces at its boundary. The strength of the restraining forces reflects the relation between the perceived nature of the barrier region and the perceived ability of the individual. The same social or physical obstacle corresponds, therefore, to different restraining forces for different individuals. It is an important characteristic of the force field, which corresponds to a barrier, that it does not extend, or at least does not extend to any important degree, outside the region of the barrier. In other words, there are only forces at the barrier away from the barrier but not forces in any other region which have the direction away from the barrier.

We have employed the term "valence" in our discussion of tension systems and of driving forces; the three constructs are intimately interrelated. A region within the life space of an individual which attracts or repulses is considered to have *valence:* a region of positive valence attracts, a region of negative valence repels. Technically, a positive valence corresponds to a force field where all forces are directed toward the same region — the region of positive valence. A negative valence corresponds to a force field where all forces have the direction away from the same region — the region of negative valence.

The discussion of the relation among tension, valence, and force may be clarified by comparing two types of theories. The first type attempts to explain behavior as a result of the inner state of the person, including the inner stimuli (which in turn may be due to stimuli from outside). This view is characteristic of many of the so-called physiological theories. It involves a derivation of behavior including directed locomotions from the state of inner personal regions. Lewin, advancing the second type of theory, holds

> . . . that such an explanation of behavior is not possible . . .
>
> (1) Forces for locomotions cannot be derived logically from tension in the person.
>
> (2) There exists a relation between certain tensions of the inner personal regions of the person and certain valences in the environment: as the tension increases, the strength of the valence tends to increase. However, the existence and the strength of the valence does not depend only upon the tension of the person but also upon certain nonpsychological "alien" factors . . . More specifically, the valence *Va* (*G*) which an object or activity *G* possesses for a person at a given time depends upon the character and state of the person *P*, and upon the perceived nature of the object or activity *G*.
>
> (3) In the same indirect way tension and forces for locomotion are related . . . the force is a function of the [strength of the] valence (*G*) and of the relative position of *P* and *G* [the strength of

> the force increases as the strength of the valence increases and as the distance between the person and the valent region decreases] . . . In other words, the force for a locomotion depends on the need or the tension of the person, on those non-psychological factors which affect the existence of a valence, and on the relative position of the person and the valence.
>
> . . . *Instead of linking the need directly to the motoric, the need is linked with certain properties of the [psychological] environment. The [psychological] environment then determines the motoric.* (Lewin, 1938, pp. 106-108)

The dynamic concepts which Lewin has utilized to characterize the psychological environment have been brilliantly employed, in conjunction with his topological concepts, to give insight into the nature of different types of *conflict situations.* Lewin's article, *The Psychological Situations of Reward and Punishment* (Lewin, 1935b, Ch. IV) should be read in its entirety to obtain the full richness of his important discussion. Here we can give only the essential points of his distinctions.

Lewin first distinguishes between the *interest* and the *punishment* situation. He points out that one can represent the fact that a person is interested in a task by a positive valence or a field of forces which all have the direction toward the goal. Hence, if difficulties force the person out of his original direction to the goal, a new vector in the direction of the goal comes into play and initiates behavior toward the goal. The behavior thus makes a pronounced goal-seeking impression. In the punishment situation, the individual finds himself between two negative valences arising from the unpleasant task and the threat of punishment for failure to do the task. If the threat of punishment is to be effective, the forces emanating from it must not only be strong enough to overcome the forces away from the unpleasant task but they must be so *distributed* as to continuously hold the person within the field of the task. When the forces emanating from the threat of punishment do not have this distribution or if they are not continuously active, the occurrence of a difficulty in pursuing the unpleasant task (unlike the case in the interest situation) will tend to push the individual in a direction opposite to that of the task to be completed. The punishment situation lacks an educationally important property that the interest situation possesses: when the child is thrust away by difficulties he does not tend to *resume of himself* the direction of the task.

Lewin pointed out *three fundamental types of conflict situation,* the punishment situation being an example of one of the major types. A conflict is characterized as a situation in which oppositely directed, simultaneously acting forces of about equal strength play upon the individual. The three fundamental types of conflict are:

1. The individual stands midway *between two positive valences* of approximately equal strength. A classical instance of this sort is that of Buridan's ass starving between two stacks of hay. However, the classical instance is highly unlikely, since this type of conflict is a condition of a labile or unstable equilibrium. As the individual, due to the play of chance factors, moves from the point of equilibrium toward one rather than the other goal region the resultant force toward that region increases and, hence, he will continue to move toward that region (and, therefore, away from the point of equilibrium). This follows from the fact that the strength of force toward a goal region decreases with increasing distance from the goal. Most instances of apparent conflict between two positive valences (e.g., between going to the theater and going to the opera) are actually conflict between regions each of which is simultaneously positive and negative in valence. One wants to go to the opera but one does not want to miss going to the theater, or one wants to go to the theater but one does not want to miss the opera — doing one precludes doing the other.

2. The second fundamental type of conflict situation occurs when the individual finds himself *between two approximately equal negative valences.* The punishment situation is an example. This type of conflict situation is very much influenced by the structure of the situation. Let us illustrate three types of conflict between negative valences: (a) A situation where the individual is between two negative valences but there are no restraints keeping him in the situation — e.g., a girl who will have to marry an unpleasant suitor or become an impoverished spinster if she remains in her village — there is nothing to prevent her from leaving her village; (b) a situation where the individual is between two negative valences but he cannot leave the field — e.g., a group member who is faced with the prospect of losing social status or of performing an unpleasant task — his situation is such that it is impossible for him to leave the group; (c) a situation where the individual is in a region of negative valence and can only leave it by going through another region of negative valence — e.g., a man is in the negative region of being cited for contempt of Congress for not testifying whether or not some of his acquaintances were members of the Com-

munist party — to purge himself of contempt he must engage in the unpleasant activity of being an "informer."

It is evident that the situation depicted in (a) will lead to a going "out of the field." Escape behavior will follow from the fact that the resultant of the forces away from the two regions of negative valence will tend to push the person in a direction that is perpendicular to the path connecting the two negative regions. Only if restraints prevent the individual from leaving the field will such a situation result in more than momentary conflict. Restraints as in (b) introduce a conflict between the driving forces related to the negative valence and the restraining forces related to the restraining barrier. There is a tendency for a barrier to acquire a negative valence which increases with the number of unsuccessful attempts to cross it and which, finally, is sufficiently strong to prevent the individual from approaching it (Fajans, 1933a, b). Thus, the conflict between driving and restraining force is replaced by a conflict between driving forces as in (c) above. This fact is particularly important for social psychology since, in many situations of life, the barriers are *social.* When a person turns against the barrier, he is in effect directing himself against the will and power of the person (s) or social group to whom the erection of the barrier is due. That is, when there is no way to escape the conflict between negative valences except by overcoming social barriers, the barriers and the people from whom he perceives that they derive will take on negative valence as the individual is unable to escape. One of the effects of the threat of punishment is to create a situation in which the individual and the creator of the barrier stand against each other as enemies.

3. The third fundamental type of conflict situation occurs when the individual is exposed to *opposing forces deriving from a positive and negative valence.* One can distinguish at least three different forms of this conflict: (a) The situation in which a region has both positive and negative valences (the Freudian concept of ambivalence is subsumable under this variety of conflict), e.g., a person wishes to join a social group but he fears that being a member will be too expensive. (b) The situation in which a person is encircled by (but is not actually in) a negative or a barrier region and is attracted to a goal which is outside the negative or barrier region — e.g., a person who has to go through the unpleasant ordeal of leaving his home or his group in order to pursue some desired activity. The nature of this situation is such that the region of the person's present activity tends to acquire negative valence as long as the region which encircles the person hinders locomotion toward desired outer goals. Thus, being a member of a minority group or being in a ghetto or in a prison often takes on negative valence, apart from the inherent characteristics of the region, because one can get to desired goals from the region only by passing through an encircling region of negative valence. (c) The situation in which a region of positive valence is encircled by or is accessible only through a region of negative valence. This type of situation differs from (b) in that the region of positive valence rather than the region in which the person is to be found is encircled by the negative valence. The "reward situation," in which the individual is granted a reward only if he performs a task (the task is unpleasant), is an example of this type of conflict. Similarly, the initiation rites in many social groups — e.g., the hazing procedures of college fraternities — may be viewed as exemplifying this kind of conflict.

Lewin (1951b, p. 259, or 1946a), as well as Miller (1944), has pointed out that the forces corresponding to a negative valence tend to decrease more rapidly as a function of psychological distance than do the forces corresponding to a positive valence. The amount of decrease depends also upon the nature of the region which has a positive or negative valence. It is different, for example, in the case of a dangerous animal which can move about than it is in the case of an immovable unpleasant object. The difference in the gradients of decrease for forces deriving from positive and negative valences accounts for the fact that a strong fear or a strong tendency to withdraw may be taken as evidence of a strong desire for the goal despite the apparent paradox. Only with a highly positive goal will the equilibrium point between approach and avoidance tendencies be close enough to the negatively valent region to produce a strong force away from the goal. On the other hand, strengthening the negative valence of a region may very well have the effect of weakening the forces in conflict, since the equilibrium point (i.e., the point where the force away from the negative valence and the force toward the positive valence are of equal strength and opposite direction) may be pushed a considerable distance away from the valent regions.

In our earlier discussion of overlapping situations, we pointed out that one could analyze conflict in terms of *overlapping situations.* The usefulness of characterizing conflict in terms of overlapping situations becomes evident when one considers more broadly *choice* situations or *decision-making* situations (of which conflict might be considered a subtype). Here it becomes apparent that the *relative potency* of each of the overlapping situations, as well as the strength of the conflicting valences and the relative position of the person in relation to the valent regions, is an important factor determining choice. Potency may be viewed as a factor influencing the strength of the effective valence or effective force. Even if a goal is highly attractive it will not elicit behavior if the potency of the situation to which the goal is linked is very low. (The combined potency of all the situations in which an individual is simultaneously involved is considered to be equal to 1; any one situation will vary between 0 and 1.) Escalona (1940) has indicated that one may consider the potency of a situation as varying directly with the perceived probability of success in obtaining the goal. While it is clear that factors other than the subjective probability of success may influence "potency," these have never been adequately discussed in the writings of Lewin or of his students.

"Reaching a decision" when there are equally attractive and opposing goals may be thought of as a process which raises one of the alternative paths to a stable, dominant potency. Experimental results have indicated that, other things being equal: the decision time is greater in choice between two negatives as compared with two positive valences (Barker, 1942), and that it increases the more the opposing situations are of equal potency (Escalona, 1940), the more the opposing forces are of equal strength (Cartwright, 1941), the stronger the opposing forces (Barker, 1942), and as a function of certain personality characteristics. Cartwright and Festinger (1943), by conceiving the region of decision between two positive valences to be bound by restraining forces against leaving the decision area and by assuming that the probability of choosing one alternative or another is distributed in the form of the normal curve, have developed a rather neat quantitative theory of decision from which many of the results of classical psychophysical studies of choice can be derived.

Let us sum up our discussion of the dynamic features of the psychological environment by pointing out that Lewin employs three main concepts — valence, driving force, and restraining force — supplemented by the concept of potency. These concepts are dealt with most fully in his monograph, *The Conceptual Representation and the Measurement of Psychological Forces.* In this monograph, he also considers, as its title suggests, the problem of the measurement of psychological forces. He points out that the strength of a force can be measured either by (1) the strength of opposed driving or restraining forces, (2) by the relative persistence of directed activity, and (3) by the velocity of locomotion or of restructuring. Lewin's monograph should be consulted for a detailed discussion of these various methods for ascertaining the strength of a force.

CONCEPTS DEALING WITH CHANGE IN THE PSYCHOLOGICAL ENVIRONMENT

Although little of Lewin's work deals with "learning" as traditionally conceived, a good deal of the research by Lewin and his students has relevance to "learning" more broadly conceived as any change in the person which produces a change in his psychological environment. Lewin (1951b, p. 67, or 1942) writes:

> Within what is called learning, we have to distinguish at least the following types of changes: (1) learning as a change in cognitive structure (knowledge), (2) learning as a change in motivation (learning to like or dislike), (3) learning as a change in group belongingness or ideology (this is an important aspect of growing into a culture), (4) learning in the meaning of voluntary control of the body musculature.

It is beyond the purpose of this chapter to discuss in any detail the various types of changes which might be called "learning." Lewin's writings on learning were largely confined to a study of "learning" as a change in valences and values, and here we shall focus our attention on this area. However, insofar as Lewin concerned himself with problems of change in cognitive structure, his viewpoint was sympathetic to Tolman's approach to learning (White, 1943). He emphasized that both differentiation and restructuring (in the sense of separating certain regions which have been connected and connecting regions which have been separated) are key processes in cognitive change. Further, he stressed that changes in the cognitive structure may occur in any part of the individual's life space, including the psychological future, the psychological present, or the psychological past. It may occur on the reality level or on the irreality level of each of these sections of the life space.

In preceding sections of this chapter, we have discussed many studies bearing on factors affecting changes in needs and valences — e.g., the studies on satiation and substitution. Here we shall confine our discussion to three topics: level of aspiration, socially induced changes, and social perception.

Level of aspiration. Perhaps no other area of research that Lewin and his students have opened to experimental investigation has been the subject of so many studies as that of *level of aspiration.* The level of aspiration may be defined as the degree of difficulty of the goal toward which the person is striving. The concept of level of aspiration is relevant only when there is a perceived range of difficulty in the attainment of possible goals and there is variation in valence among the goals along the range of difficulty.

In discussing the level of aspiration, it may be helpful to consider a sequence of events which is typical for many of the experimental studies in this area: (1) A subject plays a game (or performs a task) in which he can obtain a score (e.g., throwing darts at a target); (2) after playing the game and obtaining a given score, he is asked to tell what score he will undertake to make the next time he plays; (3) he then plays the game again and achieves another score; (4) he reacts to his second performance with feelings of success or failure, with a continuing or new level of aspiration, etc. In the foregoing sequence, point (2) (setting of the level of aspiration) and point (4) (reaction to achievement) are particularly significant for the dynamics of the level of aspiration.

In outline form, the theory of the level of aspiration is rather simple (Lewin, Dembo, Festinger, and Sears, 1944). It states that the resultant valence of any level of difficulty will be equal to the sum of the valence of achieving success times the subjective probability of success minus the valence of failure times the subjective probability of failure. The level of aspiration (i.e., the goal an individual will undertake to achieve) will be the level of difficulty that has the highest resultant positive valence. The subjective experience of success or failure is determined by the relation of the individual's performance to his level of aspiration (providing, of course, that the performance is seen to be self-accomplished) and not simply by his absolute accomplishments.

Experimental work on the level of aspiration has brought out the variety of influences which affect the positive and negative valences of different levels of difficulty. It has indicated that cultural and group factors establish *scales of reference* which help to determine the relative attractiveness of different points along a difficulty continuum. Some of these influences are rather stable and permanent in their effects. It has been found, for example, that most people of western culture, under the pervasive cultural pressures toward "self-improvement," when first exposed to a level of aspiration situation give initially a level of aspiration which is above the previous performance score, and under most conditions tend to keep their level of aspiration higher than their previous performance. In addition to broad cultural factors, the individual's level of aspiration in a task is likely to be very much influenced by the standards of the groups to which he belongs (Anderson and Brandt, 1939; Hilgard, Sait, and Magaret, 1940). The nature of the scales of reference set up by different group standards may vary.

> In some cases, for instance, in case of the ideology underlying the college term "Gentleman *C*," the group standard is equivalent to the maximum valence on the scale of success: to be either above or below this standard is considered less desirable than the standard. The fashion, particularly in democratic countries, frequently follows a similar pattern of an optimum rather than a maximum of elegance as the most desirable level. In other cases, the group standard merely indicates a level at which the valence gradient is particularly steep: there is little success valence and much negative valence of failure immediately below the group standard, and much success and little failure valence directly above group standard. (Lewin, Dembo, Festinger, Sears, p. 368)

Reference scales do not come only from membership in a definitely structured social group, for they also may reflect the influence of one's self-image, other individuals, or groups that either establish certain standards for performance or that serve as models for evaluating self-performance. Thus, the level of aspiration of a college student with respect to an intellectual task will vary depending upon whether he is told that a given score was obtained by the average high school student, the average college student, or the average graduate student (Festinger, 1942a, b).

Research has given some insight into the factors determining the values on the scale of subjective probability. A main factor which determines the subjective probability of future success and failure is the past experience of the individual in regard to his ability to reach certain objectives (Jucknat, 1937). If the individual has had considerable experience with a given activity he will know pretty well what level he can expect to reach or not to reach and the gradient of values on the subjective probability scale will be steep. However, it is not only the average of past performances which determines an individual's subjective probability scale, but also the *trend* — whether he is improving, getting worse, or remaining the same. Furthermore, there is experimental evidence to indicate that the last or most recent success or failure has a particularly great influence on the individual's expectation of his future achievement level. In addition, there is evidence that the subjective probability scales as well as the performance of others can influence the subject's own probability scale. Personality factors — e.g., self-confidence — may also influence subjective probability.

The level of aspiration theory has widespread implications for many social phenomena. It gives insight into the reasons for social apathy in the face of pressing political and international problems. People are not likely to attempt to achieve even highly valued objectives when they see no way of attaining them. Similarly, it sheds some light upon why social revolution tends to occur only after there has been a slight improvement in the situation of the oppressed groups — the improvement raises their level of aspiration, making goals which were once viewed as unattainable now perceived as realistic possibilities.

The level of aspiration technique suggests itself as an instrument to compare different cultures and to characterize their system of values in a systematic way. Such questions as optimism-pessimism, the need for success, the fear of failure, the influence of specific group standards, etc., seem open to the use of this type of technique.

Socially induced changes. Lewin makes a distinction between changes due to forces which are "imposed" upon the individual and changes which reflect directly the individual's own needs. In this connection, it is relevant to introduce the concept of *powerfield* which, although not well defined, has been employed in many social psychological investigations (Wiehe, 1935; Arsenian, 1943; Frank, 1944; Lippitt, 1940; Cartwright, 1950). A powerfield is an inducing field; it can induce changes in the life space within its area of influence. The induced changes may be with respect to the valence of regions in the life space or they may be cognitive in nature.

The source of a powerfield is usually but not necessarily a person. It is possible to speak of a person's area of direct social influence as his powerfield and to speak of the changes which can be or are induced in the life spaces of other people by this person's powerfield. For example, in a situation where workers have no intrinsic interest in their work and do not view it as a means to some important goal, they would be inclined to loaf were it not for the presence of a supervisor who can induce forces in the direction of performing work. When the supervisor is not psychologically present and when therefore his powerfield is absent, the workers will resume their loafing.

The distinction between *own* and *induced* forces has been found to be useful in explaining some of the differences in behavior under autocratic and democratic leadership (Lippitt and White, 1943). Thus, children in a club when led by authoritarian leaders (who determined policy, dictated activities, and were arbitrary and personal in evaluation of activities) tended to develop little of their own motivation with respect to club activities. Although the children worked productively when the leader was present (i.e., when his powerfield was psychologically effective), the lack of personal motivation toward group goals clearly evidenced itself in (1) change of behavior when the leader left the club, (2) absence of motivation when the leader arrived late, (3) lack of carefulness in the work, (4) lack of initiative in offering spontaneous suggestions in regard to club projects, (5) lack of pride in the products of club effort.

The distinction between own and induced forces has also been employed to give insight into why workers are usually happier and more productive when they can *participate* in the decisions which affect their work. Apart from other considerations, participation in goal setting is more likely to create own forces toward the goal, and thus there will be no necessity to exert continuous social influence to induce forces in the direction of the work goal or to induce restraining forces to prevent the individual from leaving the work region (Coch and French, 1948; McGregor, 1944).

A powerfield may be characterized not only in terms of its perceived source (e.g., a given person, a group, a personal value, a law) but also in terms of: the regions that the powerfield can affect in the individual's life space (e.g., a mathematician may influence your views about the soundness of a mathematical proof but not about the soundness of a political viewpoint); its strength or power as indicated in the magnitude of the changes which it can induce; the conditions under which a powerfield is effective (e.g., a child is unlikely to obey his father if his mother is telling him to do the opposite); the nature of the changes induced by the powerfield (e.g., whether barriers or valences are affected, whether the induced forces act upon the behaving self or regions of the psychological environment); whether or not being in the powerfield of another person is positively or negatively valent (e.g., does an individual like to have decisions made for him or does he like to make them himself); the perceived qualities of the source of the powerfield — friendly or hostile, personal or impersonal; the degree of correspondence or conflict between induced and own forces; the attributes of the source of the powerfield which give rise to its power (e.g., physical strength, social role, personal attractiveness).

Social influence and socially induced change have been the foci of a series of investigations conducted under the direction of Ronald Lippitt and Fritz Redl (Grosser, Polansky, and Lippitt, 1951; Polansky, *et al.,* 1949; Polansky, Lippitt, and Redl, 1950a, b; Lippitt, Polansky, and Rosen, 1952). These studies were conducted in three children's camps: a camp for emotionally disturbed boys, a camp for emotionally disturbed girls, and a camp for middle-class, nondisturbed boys. Two different types of social influence processes were investigated: *behavioral contagion,* which is defined as "the spontaneous pickup or imitation by other children of a behavior initiated by one member of the group where the initiator does not display any intention of getting the others to do what he did," and *direct influence,* which is indicated by behavior which has "the manifest objective of affecting the behavior of another member of the group." Various kinds of data were collected through sociometric questionnaires and through direct observation of the children in their activities. The data included measures of the power attributed by other children to each child, self-perception of power, of frequency of initiation of behavioral contagion and of direct influence, frequency of pickup of contagion and of recipiency of direct influence.

Some of the major results may be summarized as follows: (1) A child is more likely to "contage" from the behavior of another group member the higher the attributed power of the other group member. (2) A child is more likely to accept direct attempts to influence him the higher the attributed power of the group member trying to influence him. (3) The average child tends to initiate more deferential, approval-seeking behavior toward group members with high attributed power than toward those whose attributed power is low. (4) The more attributed power a child perceives himself to have the more likely he is to initiate direct-influence attempts. (5) The children with high attributed power make more attempts to influence others and are more successful in these attempts than the children with low power. (6) While neither "height" nor "weight" have an unequivocal relationship to attributed power, children who have the highest attributed power are also perceived to have the best fighting ability and to have superiority in "campcraft skills." (7) Children who have high attributed power are liked better and identified with more than are other group members.

Social perception. Heider (1953) has used the concept "heteronomous event" to include the phenomena referred to by the concept of powerfield but more broadly to refer to all changes whose source does not lie in the person but in the environment. He points out that the psychological significance of a heteronomous event is very much influenced by its perceived source. Thus, you are more likely to be angry when somebody steps on your toes if you see this event as having its source in the intentions of another person than if you see it as one of those things which happen in a crowded bus. The focus in the study of social perception, as Heider defines it, is not so much the study of social influences upon the perception of the physical world but rather the study of the conditions which influence our perception of other people (groups, etc.) and which determine the characteristics and relations that we attribute to them (Heider, 1944; Heider and Simmel, 1944; Heider, 1946).

Heider indicates that the attribution of responsibility to people as sources of events serves the function of providing a cognitive stability to our psychological environment. The motives and other personality characteristics of other persons provide the underlying causes which give coherence and meaning to the diverse "surface facts" with which we are confronted. The life space would be in a state of confusion if cognitive awareness and meaning were determined by the molecular events or proximal facts, e.g., the physical movements of parts of the body, without regard to the more distal, underlying causes which give rise to and determine the direction and significance of the movements. In effect, Heider is asserting that people develop their own "naïve" psychology as a means of giving meaning and cognitive organization to the surface actions of others. In this naïve or everyday psychology, there are concepts which deal with *perception, action, wanting, trying, being able to, experiencing* or *suffering, ought* and *may, attitudes, belonging.* Much of the task in the study of social perceptions may be defined as that of determining the conditions which govern the development and application of these concepts in our perception of other people.

Heider indicates that the attribution of responsibility for an event to another always implies that the person and event are seen as "belonging" together in a unit. In the most primitive, global form of attribution, a person is seen as the source of an event merely when the event can in any way be seen as being connected with him. With increasing sophistication in the attribution of responsibility the meaning of an event and person "belonging" together is determined by conceptions of the person's abilities and his wants or intentions. Thus, a more sophisticated form is to attribute an event to the person if he were a necessary condition for the happening even if he could not have foreseen the outcome. A still more sophisticated form is attribution to a person of any effect for which he was a necessary condition, whether intended or unintended, as long as he could have foreseen the event. A more differentiated attribution occurs when the event is attributed to a person only if it was intended by the person. The most differentiated form occurs when there is perception that even the intentions of the person are dependent partly on his environment.

The nature of attribution which occurs in any particular instance will be influenced by the need to prevent cognitive imbalance. Cognitive stability requires a congruence among causal expectations with respect to related objects. That is, for a state of complete cognitive harmony to exist, the various implications of a person's expectations or judgments of any aspect of the cognized environment cannot contradict the implications of his expectations or judgments of any other aspect of the cognized environment. Thus, if a person judges *X* to be of potential benefit to his welfare, he cannot at the same time judge that *Y* (which is also judged to be of benefit to his welfare) and *X* are antagonistic and still have a stable or balanced cognitive structure. (Let *X* and *Y* be things, people, the products of people, or the characteristics of people.) When the cognitive structure is in a state of imbalance or is threatened by imbalance, forces will arise to produce a tendency toward locomotion so as to change the psychological environment or to produce a tendency toward change in the cognition of the environment. Under conditions which do not permit locomotion, the tendency for cognitive change is enhanced.

In the preceding example of *X* and *Y*, cognitive change could occur by changing one's judgment of *X* or of *Y*, or of the relationship between *X* and *Y*, or by differentiating *X* or *Y* into independent or segregated subparts. Thus, instead of continuing to like both *X* and *Y*, who hate each other, one may come to dislike either *X* or *Y*, or to feel that they don't hate each other, or to feel that I like *X* because she is beautiful and that *Y* hates *X* because she is wealthy, and that there is no interaction directly or indirectly between the two characteristics of *X* — beauty and wealth. In general, the nature of the cognitive changes resulting from an imbalance will tend to produce the most congruence and least changes in the perceptual-cognitive field. For a discussion of the Gestalt principles determining cognitive organization and reorganization and some of the implications of these principles for social perception see Krech and Crutchfield (1948) and Asch (1952).

From assumptions which are similar to those outlined in the preceding paragraphs, and with three hypotheses about the conditions of cognitive imbalance, Heider has been able to develop a keen insight into some of the conditions which determine our perceptions in interpersonal situations. He indicates that: (a) In regard to attitudes directed toward the same entity, a balanced state exists if positive (or negative) attitudes go together; a tendency exists to see a person as being positive or negative in all respects. (b) In regard to attitudes toward an entity combined with belongingness, a balanced state exists if a person is united with the entities he likes and if he likes the entities he is united with; the converse is true for negative attitudes. (c) If two entities are seen as parts of a unit, a balanced state will exist if the parts are seen to have the same dynamic character (positive or negative). If the two entities have different dynamic characters, a balanced state can exist if they are seen to be segregated.

One may consider the above hypotheses to be true for various types of positive attitudes — e.g., to like, to love, to esteem, to value — and for various types of negative attitudes. Similarly, the statements about belonging or being part of a unit are meant to refer to different types of units — e.g., membership, belonging, causality, possession, similarity, or proximity. However, it should be emphasized that the unbalanced cognitive situation produces only a *tendency* to change; whether or not locomotion or cognitive change will actually occur depends upon the strength of the other forces which are also at work in the situation.

The reader may demonstrate for himself some of the implications of Heider's hypotheses by considering three interrelated entities — one's self (p), another (o), and a third object (x). Specify the perceived attitude (positive or negative) between any two of the entities and you can determine what the perceived relationship is likely to be between the other two. Thus, if p likes x, and p dislikes o, then one could predict a tendency to perceive o as disliking x; if p likes x and also likes o, one can predict a tendency to perceive o as liking x. The reader may introduce some additional factors by considering the relation of belongingness. For example, given that p likes o, and that o caused x, one could predict a tendency for p to like x, and so forth. The reader might find it of particular interest to consider x as an aspect of the self and to see some of the predictions one would make, from Heider's hypotheses, about attitudes toward aspects of the self under varying conditions.

We do not have the space to elaborate on the many insights which Heider's approach gives to social and self-perception and to the psychology of interpersonal relations. However, let us point out that the nature of "moral" or "ought" forces and of various socio-psychological emotions — e.g., envy, jealousy, revenge, altruism, benefiting, harming — are clarified by Heider's formulations.

While as yet little research has been conducted to test the various implications of Heider's theory, several studies (Heider and Simmel, 1944; Horowitz, Lyons, and Perlmutter, 1951; Jordan, 1953) are of relevance and provide supporting evidence. The experiment by Heider and Simmel demonstrated the significance of spatial arrangements and of kinetic structures in determining the qualities and relationships we attribute to social objects. Their experiment gives some insight into the spatial and kinetic factors making for the attribution of relationship of belonging, of "initiating" versus "receiving" an event, etc. The investigation by Horowitz, Lyons, and Perlmutter tested some of the implications of Heider's hypothesis in a study of a discussion group and obtained data which support Heider's theory. Jordan, utilizing a questionnaire, had students rate 64 situations (each of which was composed of specified relationships among p, o, and x) for their degree of pleasantness-unpleasantness. His data, although suggesting some refinements in Heider's theory, provide support for Heider's notion that a situation of cognitive imbalance produces a state of tension. An ingenious experiment by Pepitone (1950), although not conducted in the framework of Heider's theory, is relevant to some of the more complex interpersonal perceptions dealt with by Heider. This experiment indicates how the perception of another person is influenced not only by the other person's attitude toward the perceiver but also by the other person's power to benefit or harm and by the perceiver's strength of motivation to receive the benefits which the other person can bestow.

ACTION RESEARCH AND GROUP DYNAMICS

Lewin's work on the study of motivation from the first heightened his awareness of the importance of social factors for practically all aspects of psychological life. Experimental studies on such topics as goal setting, conflict, level of aspiration, play, and recall and resumption of activities clearly revealed the effects of social influences. In his first explorations of social psychology, Lewin focused his attention upon the individual in his social environment. He analyzed the social influences playing upon the child (Lewin, 1931), the differences in national character between Americans and Germans (Lewin, 1936b, or 1948) the socio-psychological problems of being a member of a minority group (Lewin, 1935a, or 1948) and the situation of the adolescent (Lewin, 1939a, or 1951b). These diverse studies stimulated an interest in the study of the conditions which affect group life. This interest saw its first fruits in the group atmosphere studies conducted at the University of Iowa and began to fructify in earnest with the establishment of the Research Center for Group Dynamics at the Massachusetts Institute of Technology in 1945.

Lewin's work on social psychology reflected two basic personal orientations which Lippitt (1947) has summed up in the term "scientific citizen." *Action research* and *group dynamics*, two terms which are closely associated with the work of Kurt Lewin in social psychology, in a sense reflect each of Lewin's personal orientations.

ACTION RESEARCH

As a citizen deeply sensitive to the world in which he lived, aware of the personal tragedies caused by oppression and prejudice, Lewin devoted much of his scientific work to furthering the understanding of the practical day-by-day problems of modern society. His book, *Resolving Social Conflicts*, presents many of his socio-psychological analyses which deal with such important social problems as the effects of prejudice, methods of facing oppression, conflict in industry, conflict in marriage, morale in time of war, changing prejudiced groups. Lewin felt that the social scientist could not only contribute to the solution of social problems but also that the study of attempts to produce change in social conditions would enable scientific insight into social processes which might not otherwise be attainable. These feelings led Lewin to help stimulate an interest in action research among various social agencies concerned with reducing prejudice.

"Action research" is a term with many meanings (Chein, Cook, Harding, 1948) but let us here elaborate only on the two major complementary values which Lewin felt might be advanced by action research. First of all, Lewin (1946a, p. 34, or 1948) pointed out that many social agencies and civic groups concerned with eliminating and preventing community problems are inffective, despite their efforts and good will. "They feel in a fog on three counts: 1. What is the present situation? 2. What are the dangers? 3. And most important of all, what shall we do?"

One of the consequences of this unclearness is the lack of standards by which to measure progress. As Lewin (1946a, p. 35) points out:

> This lack of objective standards of achievement has two severe effects:
>
> 1. It deprives the workers in intergroup relations of their legitimate desire for satisfaction on a realistic basis. Under these circumstances, satisfaction or dissatisfaction with his own achievement becomes mainly a question of temperament.
>
> 2. In a field that lacks objective standards of achievement, no learning can take place. If we cannot judge whether an action has led forward or backward, if we have no criteria for evaluating the relation between effort and achievement, there is nothing to prevent us from making the wrong conclusions and to encourage the wrong work habits. Realistic fact-finding and evaluation is a prerequisite for any learning.

In addition to the value that action research might have for social agencies, Lewin felt that the linking of research to social action might give the social scientist access to basic social processes which otherwise he would be unable to study. Further, Lewin's orientation to dynamics in individual and group psychology led him to the conclusion that change studies were necessary to reveal underlying processes. Since social scientists are rarely in the position to create social change of their own initiative, they have much to gain through cooperation with social agencies that attempt to produce social and community changes. Through such cooperation, the social scientist is enabled to do research in the community and he escapes the limitations of social psychological research on the much abused "college sophomore." Considerable experience with action research, however, has indicated that the goals of action and the goals of research may often be incompatible. The danger that confronts the research worker, in such situations, is the possibility that his research design or methodology will be sacrificed to the achievement of the social action objective.

By working closely with various social agencies, educational institutions, and industries Lewin was able to stimulate a number of field experiments on such important topics as leadership training (Bavelas, 1942), intergroup relations (Lippitt, 1949), and group productivity (Coch and French, 1948). Lewin's emphasis on field experiments in diversified community settings, his stimulation of social psychologists to leave the sheltered comfort of the college campus and to carry their experimental procedures into settlement houses, workshops, offices, and factories, are not the least of his contributions to this field.

GROUP DYNAMICS

While Lewin stressed the importance of relating action and research, he was perhaps even more deeply concerned with the relationship between research and theory. Since the 1930's, and stimulated by World War II, developments in social psychology, sociology, and cultural anthropology have brought these apparently separate disciplines into closer linkage. Lewin (1947a, b) felt that a fruitful integration of the social sciences could occur, however, only by moving from the description of social and group life to a theoretical analysis and experimental study of the dynamic problems of changing group life. This theoretical analysis and experimental study he referred to by the title *group dynamics.*

It is well to realize that Lewin's first writings in the area of group dynamics (Lewin, 1939a, b) occurred at a time when psychologists commonly denied the existence or reality of "groups." Only "individuals" were real, and to refer to characteristics of groups — e.g., "group atmosphere," "group goals," etc. — was viewed as being "nonscientific" or "mystical." One of Lewin's major contributions was to help make the concept of *group* acceptable to psychologists, that is, to lead psychologists to accept the notion that groups, *per se,* have characteristics, and that the behavior of an individual is greatly influenced by the various groups to which he belongs.

Among psychologists, the taboo against believing in the reality of the group was mainly shattered by demonstrating that "something could be done with" groups; that is, by handling groups experimentally (Lewin, 1948). The concept of group is, of course, not new; it has been the key concept in sociological writings for decades. Nevertheless, by performing experiments with groups, Lewin performed for psychologists the service of taking this concept from the plane of irreality to the plane of reality. For sociologists, the introduction of experimental studies of groups also served a radical function by opening up many of the theoretical writings of sociologists to the possibility of experimental testing. Lewin and his students have helped to demonstrate that experimental studies of small groups could shed light on large-scale social processes. The early experiments of Lippitt (1940), Lewin, Lippitt, and White (1939), French (1944), and Festinger (1947) were instrumental in initiating and giving momentum to the experimental investigations of group life, which are now accepted as a commonplace in psychology and sociology.

Concepts of group dynamics. Apart from papers dealing with group decision and social change (Lewin, 1947a and b or c and 1948), Lewin actually wrote very little on the theory of group dynamics. However, from the research investigations of his colleagues and students at the Research Center for Group Dynamics a formidable array of concepts has been gradually emerging. Such people as Bavelas, Cartwright, Festinger, French, and Lippitt have made leading contributions. (The discussion of concepts in group dynamics that follows has been considerably stimulated by an unpublished document of Dorwin Cartwright's.)

These concepts have been largely developed in isolation from the body of sociological literature which deals with related topics. Nevertheless, the sociologically oriented reader will be able to detect many parallels between the concepts of group dynamics and of sociology. Whether or not the isolation from prior sociological theorizing has been sensible is hard to judge. The group dynamicist might argue that the isolation has produced concepts which have stimulated experimental research studies — studies which did not flow from the more general sociological conceptions. The sociologist might argue that the studies of group dynamicists might have been more sophisticated if they had been acquainted with the prior, relevant writings of sociologists.

Let us begin our discussion of the concepts of group dynamics by briefly considering the concept of *group*. Lewin (1948, p. 54, or 1943d) wrote that:

> The essence of a group is not the similarity or dissimilarity of its members, but their interdependence. A group can be characterized as a "dynamical whole"; this means that a change in the state of any subpart changes the state of any other subpart. The degree of interdependence of the subparts of members of the group varies all the way from a "loose" mass to a compact unit.

French (1944) pointed out that in addition to interdependence, membership in a group presupposes identification with the group. Deutsch (1949a) indicated that the interdependence is that of *promotive or cooperative interdependence* rather than, for example, competitive interdependence. Thus, a group may be defined tentatively as being composed of a set of members who mutually perceive themselves to be cooperatively or promotively interdependent in some respect (s) and to some degree. Our discussion of groups will throughout refer to psychological groups rather than to collections of individuals who may be objectively interdependent without, at the same time, being subjectively or psychologically interdependent.

A group may be thought of in terms of its external relations (i.e., its relationship to its external environment) in a manner analogous to that of the behaving self in the psychological environment, or it may be considered in terms of its internal relations (i.e., the relations among the members composing it) in a manner analogous to the discussion of the person and the tension systems composing it. In group dynamics, there has been very little discussion of the external relations of groups, of "group spaces." This is, of course, a profound neglect, comparable to the earlier tendency in psychology to consider the individual without relation to his environment. Most of the conceptualizations and research in group dynamics have been on the internal characteristics of groups — on their internal structural and dynamic properties.

One of the key concepts, which has been the subject of much experimental investigation, is that of *cohesiveness*. Intuitively, cohesiveness refers to the forces which bind the parts of a group together and which, thus, resist disruptive influences. Hence, the study of conditions affecting group cohesiveness and of the effects upon group functioning of variations in group cohesiveness is at the heart of the study of group life.

Festinger, Schachter, and Back (1950) have defined cohesiveness, in terms of the group member, as "the total field of forces which act on members to remain in the group." The nature and strength of the forces acting upon a member to remain in the group may vary from member to member. These investigators make the suggestion that group cohesiveness "be related to the average magnitude of this force in all parts of the group." Deutsch (1949a) in a definition which has essentially the same implications as that of Festinger, Schachter, and Back, relates group cohesiveness to the degree of perceived cooperative interdependence among members and to the strength of goals about which the members are cooperatively interdependent. "Membership motive" in the individual is defined so as to be the counterpart of cohesiveness in the group. In the foregoing definitions, the relationship between the *group* attribute of cohesiveness and the *individual* attribute of membership motive (or cohesiveness) is not adequately handled. It is obvious, however, that the average of the various individuals' membership motives is an inadequate measure of group cohesiveness. Intuitively, group cohesiveness will be affected by the distribution of membership motives in a group — i.e., whether there is much variation or little variation in membership motive among the members and whether "important" members or "unimportant" members have high or low membership motives. Parenthetically, it should be noted that in the definitions of cohesiveness the "forces to remain in the group" are a resultant of the forces to remain in the group and of the forces to leave the group. For simplicity of presentation, the discussion of cohesiveness neglects the forces to leave the group.

The force acting upon an individual to remain in the group may derive from the attractiveness or positive valence of the group, the negative valence of not being in the group, and/or the barriers or restraints against leaving the group; the forces acting upon the individual may be induced or own forces. Little research has been done to determine the group consequences of these different types of membership motivation. Back (1951), however, has shown that it makes no difference with certain aspects of group functioning whether attractiveness to a group is based upon (1) liking the other group members, (2) liking it because it is prestigeful to belong to it, or (3) liking the group because it is a means of obtaining personal goals.

In Back's experiment, subjects (in pairs) worked cooperatively in interpreting a set of pictures. The groups of pairs were formed so as to be more or less attractive to their members; the three different bases of attraction indicated in the preceding paragraph were employed. Back found that in the more cohesive groups, the members made more effort to reach agreement and were more influenced by the discussion than were members of the lesser cohesive groups no matter what the basis of attractiveness. However, discussion was influenced by the source of attractiveness to the group. The behaviors of the people in the groups composed of people attracted to the group by a liking for the other group members, were more chatty than the other groups; by the prestige of the group, were more cautious and less related to other members; by the group as a means to a goal, were more impersonal and task-oriented.

While there have been few direct studies of the *factors determining group cohesiveness,* a number of experimental investigations incidentally bear on this area. Thus Back (1951) found that he could produce high, as contrasted with low, cohesiveness in experimentally created groups by stressing to the subjects how much they would like each other, or how important it was for the group to do well on the task since the task was a test of ability, or by stressing the prestige of the group. French (1941) found that groups composed of students who belonged to the same athletic teams were more cohesive than groups composed of students not previously acquainted. Deutsch (1949b) and Levy (1953) found that groups rewarded on a cooperative basis were more cohesive than groups rewarded on a competitive basis. Lippitt and White (1943) in a study of the effects of autocratic, democratic and *laissez faire* group leadership in boys' clubs found that democratically led groups were most cohesive. Schachter (1951) produced clubs with high cohesiveness by composing them of students who expressed moderate or high interest in their activities; he created low cohesive clubs by composing them of students who expressed little or no interest in their activities. Thibaut (1950), in the most direct study of factors affecting cohesiveness, found that being in a persistently high-status group increases cohesiveness; on the other hand, being in a persistently low-status group tends to decrease the attractiveness of one's own group but also increases the hostility toward the high-status group. Those members of the low-status group whose hostility toward the high-status group is most aroused, the most central members, may as a consequence of their rejection of the other group become more cohesive.

Various measures of "cohesiveness" have been employed in the foregoing experimental investigations: desire to remain in the group, the ratio of "we" remarks to "I" remarks during group discussion, ratings of friendliness, evaluations of the group and its product, acceptance of each other's ideas, in-group sociometric choice versus out-group sociometric choice, etc. Deutsch (1949a), in a theoretical paper, provides a rationale for the use of a wide variety of measures of membership motive by developing the hypotheses that members of more cohesive (cooperative) groups as compared with members of less cohesive (competitive) groups would, under conditions of successful locomotion: (a) be more ready to accept the actions of other group members as *substitutable* for similarly intended actions of their own (and therefore would not have to perform them also); (b) be more ready to accept *inductions* (i.e., be influenced) by other members; and (c) be more likely to *positively cathect* the actions of other group members. From these core hypotheses, with the addition of more specific assumptions, it is possible to derive the influence of more or less cohesiveness upon many aspects of group functioning. Thus, from the substitutability hypothesis, it is possible to predict that more specialization of function, more subdivision of activity, more diversity of membership behavior would occur in the more cohesive groups. The inducibility hypothesis leads to the prediction that members of more cohesive groups would be more attentive to one another, be more understood by one another, be more influenced by one another, be more likely to change and have more internalization of group norms than members of less cohesive groups. The cathexis hypothesis leads to predictions of greater friendliness, greater ratio of in-group sociometric choices, etc., in the more cohesive groups. Data in a variety of experiments (Deutsch, 1949b; Levy, 1953; Schachter, 1951; Back, 1951; Gerard, 1953) support the foregoing predictions.

In a well-integrated program of research, Festinger and his co-workers have developed a series of fertile hypotheses (Festinger, 1950) and conducted some ingenious experiments on *the communication process within groups.* In brief, these investigators have been concerned with three sources of pressures to communicate within groups: (a) communications arising from pressures toward uniformity in a group (Back, 1951; Festinger and Thibaut, 1951; Schachter, 1951); (b) communications arising from forces to locomote in a social structure (Kelley, 1951; Thibaut, 1950); and (c) communications arising from the existence of emotional states (Thibaut, 1950; Thibaut and Coules, 1952). Most of the theoretical writings and experimental work of Festinger and his associates have been concerned with (a) above.

Festinger (1950) has indicated two major sources of pressures toward uniformity in a group: social reality and group locomotion. He indicates that when there is no simple objective basis for determining the validity of one's beliefs, one is dependent upon social reality (i.e., upon the consensus of judgment among people whose judgments one will respect) to establish confidence in one's beliefs. For example, the belief that "Negroes adversely affect property values when they move into a white neighborhood" is difficult for any white property owner to test in a simple, objective way; thus, he relies upon the judgment of others as a test of its validity. Lack of agreement among members of a group provides an unstable basis for beliefs which depend upon social consensus for their support, and hence (in line with Heider's discussion of the tendency toward cognitive balance) forces will arise to produce uniformity. Pressures toward uniformity among members of a group may also arise because such uniformity is desirable or necessary in order for the group to locomote toward some goal. Greater uniformity in opinion within a group can be achieved in either of the following ways: (a) by actions (i.e., communications) that are directed at changing the positions the members occupy relative to one another — e.g., by attempting to influence others to change their opinions or by changing one's own views; or (b) by actions to make others incomparable in the sense that they are no longer effective as a comparison for one's opinions — e.g., by rejecting or excluding people with deviating opinions from the group.

Experiments have shown that increasing the attraction to the group, and thus increasing the importance of the group as a comparison object, increases the amount of influence which is attempted and the amount of opinion change which occurs when there is opinion discrepancy within a group (Back, 1951). It has also been demonstrated that there will be more pressure for uniformity in a group, the more relevant or important the opinion is for the functioning of the group (Schachter, 1951; Festinger and Thibaut, 1951). Also, as may be expected from theoretical considerations, there is evidence that when pressures toward uniformity exist, the concern is mainly with those members of the group who have opinions most divergent from one's own. Thus, members exert influence mainly on those whose opinions are most divergent from their own (Festinger and Thibaut, 1951; Schachter, 1951). While there has, as yet, been little research to indicate the conditions which will lead the pressure toward uniformity to manifest itself in a rejection of deviates rather than in an attempt to influence them to change, experiments by Gerard (1951) and by Festinger and Thibaut (1951) suggest that as the heterogeneity of a group increases, rejection of deviates also increases.

The experiments revealing the tendency to direct communication toward deviants in the group, and through communication to exert social pressure upon them, provide support for Lewin's *theory of group decision and social change*. Lewin (1947c) began his analysis of change by pointing out that the status quo in social life is not a static affair but a dynamic process that flows on but still keeps a recognizable form. He borrowed the term "quasi-stationary equilibria" from physics to apply to such on-going processes, which are kept at their present level by fields of forces preventing a rise or fall. The field of forces in the neighborhood of the level of equilibrium presupposes that the forces against going higher than the equilibrium level increase with the amount of raising and that the forces against lowering increase (or remain constant) with the lowering. Thus, if we assume that a group standard is operating to determine the level of worker productivity in a factory, any attempt upon the part of the worker to deviate from the standard by higher productivity will only result in stronger forces being induced upon him by his co-workers to push him back into line. That is, as the Festinger experiments have demonstrated, the deviant will be exposed to stronger pressures toward uniformity the more he deviates. However, as Lewin points out, the gradient of forces may change at a distance from the equilibrium level so that after an individual has gone a certain distance from the equilibrium level, the forces may push him away rather than pull him toward the group standard.

Lewin's analysis of the status quo as a quasi-stationary equilibrium has two major implications. First of all, it points out that change from the status quo can be produced either by adding forces in the desired direction or by diminishing opposing forces. However, the two methods of producing change have different consequences: by adding forces, the process on the new level would be accompanied by a state of relatively high tension, since the strength of forces in opposition would be greater; by decreasing the opposing forces, the new level would be accompanied by lower tension. Secondly, it highlights the difficulties of attempting to change individual conduct and attitudes that are rooted in groups by efforts which are directed at the individual and not his group. If one endeavors to change the prejudices of an individual without changing the prejudice of the group to which he belongs, the individual will either be estranged from his group or will be under pressure from his group to revert to his initial attitude. Isolated individuals may perhaps change their attitudes because of their individual experiences, but the person who is deeply enmeshed in the social life of his community is unlikely to be able to resist the pressures to conform on matters of community importance if he wishes to continue in good standing.

Considerations such as the foregoing have led to a series of experiments in various settings — in the school, with neighborhood groups, in industry, in an interracial workshop — on the relative efficacy of changing behavior by efforts directed at individuals or at a group (Lewin, 1948). A typical procedure has been to compare the results of a lecture or individual instruction in changing behavior with respect to the use of certain foods with the results of a group discussion on the merits of the food combined with a group decision favoring the use. Results have clearly indicated that the group decision method produces more change. However, most of the experiments on group decision were unfortunately not designed in such a way as to rule out the possibility that such factors as interest or decision, *per se,* were not the factors producing the greater change. A more carefully controlled experiment by Bennett (1952) suggests that the apparent advantages of the group-discussion method result primarily from the fact that group discussion facilitates decision and perception of consensus. She concludes that group discussion, other things being equal, was not more effective than lecturing, nor did public identification of individuals' decisions contribute appreciably to obtained differences. In Bennett's experiment, the subjects were college students in introductory psychology classes and the content of their lecture or group discussion was concerned with a request to volunteer as experimental subjects for an unspecified psychological experiment. The "groups" were not very cohesive (introductory psychology sections) nor was the issue of volunteering group-rooted in the sense of being clearly relevant to group norms. The cumulative evidence as well as everyday experience supports Lewin's basic contention that to change group-rooted individual attitudes one must, in many instances, change the group to which the individual belongs.

Horowitz (1950) performed an experiment which demonstrates some of the psychological consequences of group decision. He designed an experimental setting where it was possible for members to set up goals for the group and to perceive where the group stood in regard to the attainment of the goals. In the course of working toward the attainment of some of the group-agreed-upon goals the group was interrupted. In others, the group was allowed to complete the task. After working on a number of different tasks, a Zeigarnik-type test for memory of completed and of interrupted tasks was given to all members of the group. The experiment was conducted in such a way that each individual indicated by vote whether or not he wanted to complete each task. The results of a secret ballot for each task were also announced to the group. (The experimenter actually announced a falsified group vote in order to be able to study its effects under equated conditions.) The results clearly indicated the motivational effect of the group-agreed-upon goals: in general, more interrupted than completed tasks were recalled. In addition, there was clear evidence that the announced group vote (i.e., group decision) whether to complete or not to complete the task affected recall: when the announced group vote was to abandon the task recall was smaller than when the announced group vote was to complete the task.

Much of our discussion of group dynamics, so far, has considered the group as made up of homogeneous parts. In reality, of course, groups are differentiated structures, and the behavior of any group member will be largely determined by his position in the group structure. Although the concepts of *group structure* and of *position within the group* have not been well defined in the writings of group dynamicists, a good deal of research related to these concepts has taken place. Thus, Lewin, Lippitt, and White's (1939) pioneering study on the effects of different types of leadership behavior has demonstrated how crucial the position of leadership is in determining the atmosphere of the group. Bavelas (1951) and Leavitt (1951) have demonstrated the significance of position in a communication network in determining member behavior and have indicated how the structure of a communication network may affect group productivity and individual satisfactions. Kelley (1951) has experimentally shown the significance of position in a status hierarchy in determining the nature and direction of communication. Polansky, Lippitt, and Redl (1950), in a study of behavioral contagion in groups, have explored the relationship between status (as defined by various criteria in sociometric choice) and the ability to influence others in the group. Studies such as the foregoing provide the rationale for use of the communication process (who communicates with whom, about what, with what frequency, in what manner, in what circumstances, and with what effects) as a key instrument for characterizing group structure and for locating the occupants of various positions within this structure.

The very diversity of these studies indicates some of the varied meanings that the term "position" may have. "Position" has been employed to refer to the functions a member performs in a group, to an individual's locus in a communication network, to a person's ability to induce forces, to a person's prestige in the group, etc. Gerard and others (unpublished memorandum) provide insight into the reason for the diversity in meanings of "position." They have found it useful to conceptualize the position of a person in a group "as a locus which describes his relationship to other group members with respect to a single dimension. At any given time each member may have one and only one position on a specified dimension." A dimension refers to any group-psychological aspect which may be useful in describing the relations of one group member to another. Examples of dimensions are (a) being able to communicate with, (b) having power over, (c) having prestige for, (d) having more skill than. In some social settings one dimension may become more salient than others and thus one's locus on that dimension becomes most important for the individual. This saliency is termed *dimension dominance;* the dominant dimension is most powerful in determining a given range of behavior. Thus, to characterize adequately any group member's relations to others in the group over a period of time and in different social settings one will have to locate him along a number of dimensions, i.e., in a number of different positions.

As we indicated in our earlier discussion of psychological ecology, a position in the group may be thought of as a "behavior setting" — i.e., as a part of the nonpsychological milieu which induces forces having certain magnitudes and directions upon an individual who is located in it. Behavior settings have their coercive and restraining powers as a result of a social norm or group standard that defines the behavior setting as a locus which is appropriate for particular kinds of behavior. It follows that the ability of a behavior setting or a position in the group to induce behavior will, in part, be dependent upon the strength of membership motive in the individual.

Let me conclude the discussion of group dynamics by reminding the reader that the experimental study of small group functioning began in earnest a little over fifteen years ago. There has been an extremely rapid proliferation of studies and of theory since then. While Lewin and his students did much to give momentum to the development of "group dynamics," it is no longer their exclusive domain — it is an integral part of social psychology.

SUMMARY

In this chapter, we have attempted to present a survey of the theoretical approach to psychology which is commonly termed "field theory." While many theorists in psychology might be discussed under the title "field theory in social psychology," we have limited our discussion to the work of Kurt Lewin and the people who have been closely associated with him. We discussed Lewin's approach to theorizing in some detail because of our belief that his extraordinary creativeness as theorist and as experimenter reflected "a way of thinking." We also surveyed the basic concepts of Lewin's "topological and vector" psychology and some of the research studies which grew out of these concepts. Finally, we discussed action research and group dynamics.

The large sweep of Lewin's work, the brilliant innovations in many areas of psychology, cannot help but be impressive. Yet, like many trail blazers, the paths he opened up were left rough-hewn and unfinished. For orderly, safe passage brigades of scientific workers are required to fill in the holes and widen the paths.

REFERENCES

ACH, N. *Uber den Willensakt und das Temperament: eine experimentelle Untersuchung.* Leipzig: Quelle and Meyer, 1910.

ADLER, D. L., & KOUNIN, J. Some factors operating at the moment of resumption of interrupted tasks. *J. Psychol.,* 1939, *7,* 355-367.

ALPER, THELMA G. Memory for completed and incompleted tasks as a function of personality: an analysis of group data. *J. abnorm. soc. Psychol.,* 1946, *41,* 403-420.

ALPER, THELMA G. Memory for completed and incompleted tasks as a function of personality: correlation between experimental and personality data. *J. Pers.,* 1948, *17,* 104-137.

ANDERSON, H. H., & BRANDT, H. F. A study of motivation, involving self-announced goals of fifth-grade children and the concept of level of aspiration. *J. soc. Psychol.,* 1939, *10,* 209-232.

ANGYAL, A. *Foundations for a science of personality.* New York: The Commonwealth Fund, 1941.

ARSENIAN, JEAN M. Young children in an insecure situation. *J. abnorm. soc. Psychol.,* 1943, *38,* 225-249.

ASCH, S. E. *Social psychology.* New York: Prentice-Hall, 1952.

BACK, K. The exertion of influence through social communication. *J. abnorm. soc. Psychol.,* 1951, *46,* 9-23.

BARKER, R. G. An experimental study of the resolution of conflict by children. In Q. McNemar & K. Merrill (Eds.), *Studies in personality.* New York: McGraw-Hill, 1942. Pp. 13-34.

BARKER, R. G., DEMBO, TAMARA, & LEWIN, K. Frustration and regression: an experiment with young children. *Univ. Ia. Stud. Child Welf.,* 1941, *18,* No. 1.

BARKER, R. G., & WRIGHT, H. F. Psychological ecology and the problem of psycho-social development. *Child Devel.,* 1949, *20,* 131-144.

BARKER, R. G., WRIGHT, BEATRICE A., & GONICK, MOLLIE R. *Adjustment to physical handicap and illness: a survey of the social psychology of physique and disability.* New York: Social Science Research Council, 1946.

BAVELAS, A. Morale and the training of leaders. In G. Watson (Ed.), *Civilian morale.* Boston: Houghton Mifflin, 1942. Pp. 143-165.

BAVELAS, A. A mathematical model for group structure. *Applied Anthrop.,* 1948, *7,* 16-30.

BAVELAS, A. Communication patterns in task-oriented groups. In D. Lerner & H. D. Lasswell (Eds.), *The policy sciences; recent developments in scope and method.* Stanford: Stanford Univ. Press, 1951. Pp. 193-202.

BENNETT, EDITH B. The relationship of group discussion, decision, commitment and consensus to individual action. Unpublished doctor's dissertation, Univ. of Michigan, 1952.

BRUNSWIK, E. Organismic achievement and environmental probability. *Psychol. Rev.,* 1943, *50,* 255-272.

CARTWRIGHT, D. Decision-time in relation to the differentiation of the phenomenal field. *Psychol. Rev.,* 1941, *48,* 425-442.

CARTWRIGHT, D. Emotional dimensions of group life. In M. L. Raymert (Ed.), *Feelings and emotions.* New York: McGraw-Hill, 1950. Pp. 439-447.

CARTWRIGHT, D., & FESTINGER, L. A quantitative theory of decision. *Psychol. Rev.,* 1943, *50,* 595-621.

CASSIRER, E. *Substance and function, and Einstein's theory of relativity.* Chicago: Open Court, 1923.

CHEIN, I. The genetic factor in ahistorical psychology. *J. gen. Psychol.,* 1947, *36,* 151-172.

CHEIN, I. Environment as a determinant of behavior. *J. soc. Psychol.,* 1954, *39,* 115-127.

CHEIN, I., COOK, S. W., & HARDING, J. The field of action research. *Amer. Psychologist,* 1948, *3,* 43-50.

COCH, L., & FRENCH, J. R. P., JR. Overcoming resistance to change. *Hum. Relat.,* 1948, *1,* 512-532.

DEMBO, TAMARA Der Arger als dynamisches Problem. *Psychol. Forsch.,* 1931, *15,* 1-144.

DEUTSCH, K. Mechanism, organism, and society. *Phil. Sci.,* 1951, *18,* 230-252.

DEUTSCH, M. A theory of cooperation and competition. *Hum. Relat.,* 1949a, *2,* 129-152.

DEUTSCH, M. An experimental study of the effects of cooperation and competition upon group processes. *Hum. Relat.,* 1949b, *2,* 199-232.

EINSTEIN, A., & INFELD, L. *The evolution of physics.* New York: Simon & Schuster, 1938.

ESCALONA, SYBILLE K. The effect of success and failure upon the level of aspiration and behavior in manic-depressive psychoses. *Univ. Ia. Stud. Child Welf.,* 1940, *16,* 199-307.

FAJANS, SARA Die Bedentung der Entfernung fur die Starke eines Aufforderungscharakters beim Saugling und Kleinkind. *Psychol. Forsch.,* 1933a, *17,* 215-267.

FAJANS, SARA Erfolg, Ausdauer und Aktivitat beim Saugling und Kleinkind *Psychol. Forsch.,* 1933b, *17,* 268-305.

FESTINGER, L. Wish, expectation and group performance as factors influencing level of aspiration. *J. abnorm. soc. Psychol.,* 1942a, *37,* 184-200.

FESTINGER, L. A theoretical interpretation of shifts in level of aspiration. *Psychol. Rev.,* 1942b, *49,* 235-250.

FESTINGER, L. The role of group belongingness in a voting situation. *Hum. Relat.,* 1947, *1,* 154-181.

FESTINGER, L. Informal social communication. *Psychol. Rev.,* 1950, *57,* 271-282.

FESTINGER, L., SCHACHTER, S., & BACK, K. *Social pressures in informal groups: a study of a housing project.* New York: Harper, 1950.

FESTINGER, L., & THIBAUT, J. Interpersonal communication in small groups. *J. abnorm. soc. Psychol.,* 1951, *46,* 92-99.

FRANK, J. D. Experimental studies of personal pressure and resistance. *J. gen. Psychol.,* 1944a, *30,* 23-43.

FRANK, J. D. Experimental studies of personal pressure and resistance. *J. gen. Psychol.,* 1944b, *30,* 44-64.

FRENCH, J. R. P., JR. The disruption and cohesion of groups. *J. abnorm. soc. Psychol.,* 1941, *36,* 361-377.

FRENCH, J. R. P., JR. Organized and unorganized groups under fear and frustration. *Univ. Ia. Stud. Child Welf.,* 1944, *20,* 299-308.

GERARD, H. B. The effect of different dimensions of disagreement on the communication process in small groups. Unpublished doctor's dissertation, Univ. of Michigan, 1951.

GERARD, H. B. The anchorage of opinions in face-to-face groups. *Hum. Relat.,* in press.

GLIXMAN, A. F. Recall of completed and incompleted activities under varying degrees of stress. *J. exp. Psychol.,* 1949, *39,* 281-295.

GROSSER, D., POLANSKY, N., & LIPPITT, R. A laboratory study of behavioral contagion. *Hum. Relat.,* 1951, *4,* 115-142.

HARARY, F., & NORMAN, R. *Group structures and communication networks in terms of the mathematical theory of graphs.* Univ. of Michigan, Institute of Social Research, 1952. Pp. 1-42.

HEIDER, F. Ding und Medium. *Symposium,* 1927, *1,* 109-157.

HEIDER, F. Social perception and phenomenal causality. *Psychol. Rev.,* 1944, *51,* 358-374.

HEIDER, F. Attitudes and cognitive organization. *J. of Psychol.,* 1946, *21,* 107-112.

HEIDER, F. *Psychology of interpersonal relations.* (Tentative title.) Unpublished manuscript, 1953.

HEIDER, F., & SIMMEL, MARIANNE An experimental study of apparent movement. *Amer. J. Psychol.,* 1944, *57,* 243-259.

HENLE, MARY An experimental investigation of dynamic and structural determinants of substitution. *Contr. Psychol. Theor.,* 1942, *2,* No. 3.

HILGARD, E. R., SAIT, E. M., & MAGARET, G. A. Level of aspiration as affected by relative standing in an experimental group. *J. exp. Psychol.,* 1940, *27,* 411-421.

HOROWITZ, M. The effects of group goal-setting and locomotion on motivational processes in the individual. Unpublished doctor's dissertation, Univ. of Michigan, 1950.

HOROWITZ, M. W., LYONS, J., & PERLMUTTER, H. V. Induction of forces in discussion groups. *Hum. Relat.,* 1951, *4,* 57-76.

JAHODA, MARIE, DEUTSCH, M., & COOK, S. *Research methods in social relations.* Vol. I., New York: Dryden Press, 1951.

JORDAN, N. Behavioral forces that are a function of attitudes and of cognitive organization. *Hum. Relat.,* 1951, *6,* 273-288.

JUCKNAT, M. Leistung, Anspruchsniveau und Selbstbewusstsein. *Psychol. Forsch.,* 1937, *22,* 89-179.

KARSTEN, ANITRA Psychische Sattigung. *Psychol. Forsch.,* 1928, *10,* 142-154.

KELLEY, H. H. Communication in experimentally created hierarchies. *Hum. Relat.,* 1951, *4,* 39-56.

KOFFKA, K. *Principles of gestalt psychology.* New York: Harcourt, Brace, 1935.

KOUNIN, J. S. Experimental studies of rigidity. I *Charact. Pers.,* 1941a, *9,* 251-272.

KOUNIN, J. S. Experimental studies of rigidity. II. *Charact. Pers.,* 1941b, *9,* 273-282.

KRECH, D., & CRUTCHFIELD, R. S. *Theories and problems of social psychology.* New York: McGraw-Hill, 1948.

LEAVITT, H. Some effects of certain communication patterns on group performance. *J. abnorm. soc. Psychol.,* 1951, *46,* 38-50.

LEEPER, R. W. *Lewin's topological and vector psychology: a digest and a critique.* Eugene, Ore.: Univ. of Oregon Press, 1943.

LEVY, S. Experimental study of group norms: the effects of group cohesiveness upon social conformity. Unpublished doctor's dissertation, New York Univ., 1953.

LEWIN, K. Die psychische Tatigkeit bei der Hemmung von Willensvorgangen und das Grundgesetz der Assoziation. *Ztschr. Psychol.,* 1917, *77,* 212-247.

LEWIN, K. Das Problem der Willensmessung und das Grundgesetz der Assoziation. *Psychol. Forsch.,* 1922a, *1,* 191-302.

LEWIN, K. Das Problem der Willensmessung und das Grundgesetz der Assoziation. *Psychol. Forsch.,* 1922b; *2,* 65-140.

LEWIN, K. Environmental forces in child behavior and development. In C. Murchison (Ed.), *A handbook of child psychology.* Worcester, Mass.: Clark Univ. Press, 1931. Pp. 94-127.

†LEWIN, K. Psycho-sociological problems of a minority group. *Charact. Pers.,* 1935a, *3,* 175-187.

LEWIN, K. *A dynamic theory of personality.* New York: McGraw-Hill, 1935b.

LEWIN, K. *Principles of topological psychology.* New York: McGraw-Hill, 1936a.

†LEWIN, K. Some social psychological differences between the United States and Germany. *Charact. Pers.,* 1936b, *4,* 265-293.

LEWIN, K. The conceptual representation and measurement of psychological forces. *Contr. psychol. Theor.,* 1938, *1,* No. 4.

*LEWIN, K. Field theory and experiment in social psychology: concepts and methods. *Amer. J. Sociol.,* 1939a, *44,* 868-896.

†LEWIN, K. Experiments in social space. *Harvard Educational Rev.,* 1939b, *9,* 21-32.

*LEWIN, K. Formalization and progress in psychology. *Univ. Ia. Stud. Child Welf.,* 1940, 16, No. 3, 9-42.

*LEWIN, K. Analysis of the concepts whole, differentiation, and unity. *Univ. Ia. Stud. Child Welf.,* 1941, *18,* No. 1, 226-261.

*LEWIN, K. Field theory of learning. Yearbook, *Nat'l. Soc. Stud. Educ.,* 1942, *41,* Part II, 215-242.

*LEWIN, K. Defining the "field at a given time." *Psychol. Rev.,* 1943a, *50,* 292-310.

*LEWIN, K. Psychology and the process of group living. *J. soc. Psychol.,* 1943b, *17,* 119-129.

*LEWIN, K. Forces behind food habits and methods of change. *Bull. Nat. Res. Council,* 1943c, *108,* 35-65.

†LEWIN, K. The special case of Germany. *Publ. Opin. Quart.,* 1943d, 7, 555-566.

*LEWIN, K. Constructs in psychology and psychological ecology. *Univ. Ia. Stud. Child Welf.,* 1944, *20,* 1-29.

†LEWIN, K. Action research and minority problems. *J. soc. Issues,* 1946a, *2,* 34-64.

*LEWIN, K. Behavior and development as a function of the total situation. In L. Carmichael (Ed.), *Manual of child psychology.* New York: Wiley & Sons, 1946b. Pp. 791-844.

*LEWIN, K. Frontiers in group dynamics. I. *Hum. Relat.,* 1947a, *1,* 2-38.

*LEWIN, K. Frontiers in group dynamics. II. *Hum. Relat.,* 1947b, *1,* 143-153.

LEWIN, K. Group decision and social change. In T. Newcomb & E. Hartley (Eds.), *Readings in social psychology.* New York: Henry Holt, 1947c. Pp. 330-344.

LEWIN, K. *Resolving social conflicts.* New York: Harper, 1948.

LEWIN, K. Intention, will and need. In D. Rapaport (Ed.), *Organization and pathology of thought.* New York: Columbia Univ. Press, 1951a. Pp. 95-151.

LEWIN, K. *Field theory in social science.* New York: Harper, 1951b.

LEWIN, K., DEMBO, TAMARA, FESTINGER, L., & SEARS, PAULINE Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: The Ronald Press Co., 1944. Pp. 333-378.

LEWIN, K., LIPPITT, R., & WHITE, R. K. Patterns of aggressive behavior in experimentally created "social climates." *J. soc. Psychol.,* 1939, *10,* 271-279.

LEWIS, HELEN B. An experimental study of the role of the ego in work. I. The role of the ego in cooperative work. *J. exp. Psychol.,* 1944, *34,* 113-126.

LEWIS, HELEN B., & FRANKLIN, M. An experimental study of the role of the ego in work. II. The significance of task-orientation in work. *J. exp. Psychol.,* 1944, *31,* 195-215.

LIPPITT, R. An experimental study of authoritarian and democratic group atmospheres. Studies in topological and vector psychology. I. *Univ. Ia. Stud. Child Welf.,* 1940, *16,* No. 3, 44-195.

LIPPITT, R. Kurt Lewin, adventures in the exploration of inter-dependence. *Sociometry Monogr.,* 1947, *17,* 22-28.

LIPPITT, R. *Training in community relations: a research exploration toward new group skills.* New York: Harper, 1949.

*LIPPITT, R., POLANSKY, N., & ROSEN, S. The dynamics of power. *Hum. Relat.,* 1952, *5,* 37-64.

---

† Reprinted in *Resolving social conflicts.* (Lewin, 1948)

* Reprinted in *Field theory in social science.* (Lewin, 1951b)

LIPPITT, R., & WHITE, R. K. The "social climate" of children's groups. In R. G. Barker, J. S. Kounin, & H. F. Wright (Eds.), *Child behavior and development.* New York: McGraw-Hill, 1943. Pp. 485-508.

LISSNER, KATE Die Entspannung von Bedurfnissen durch Ersatzhandlungen. *Psychol. Forsch.,* 1933, *18,* 27-89.

LITTMAN, R. A., & ROSEN, E. Molar and molecular. *Psychol. Rev.,* 1950, *57,* 58-65.

MCGREGOR, D. Conditions of effective leadership in the industrial organization. *J. consult. Psychol.,* 1944, *8,* 55-63.

MAHLER, WERA Ersatzhandlungen verschisdemen Realitatsgrades. *Psychol. Forsch.,* 1933, *18,* 27-89.

MARROW, A. J. Goal tension and recall. I. *J. gen. Psychol.,* 1938a *19,* 3-35.

MARROW, A. J. Goal tension and recall. II. *J. gen. Psychol.,* 1938b, *19,* 37-64.

MAXWELL, C. *Matter and motion.* New York: Macmillan, 1921.

MILLER, N. E. Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: The Ronald Press, 1944. Vol. I. Pp. 431-465.

OVSIANKINA, MARIA Die Wiederaufnahne von unterbrochener Handlungen. *Psychol. Forsch.,* 1928, *11,* 302-379.

PARSONS, T., & SHILS, E. A. (Eds.), *Toward a general theory of action.* Cambridge: Harvard Univ. Press, 1951.

PEPITONE, A. D. Motivational effects in the perception of social "gatekeepers." *Hum. Relat.,* 1951, *4,* 57-76.

POLANSKY, N., FREEMAN, W., HOROWITZ, M., IRWIN, L., PAPANIA, N., RAPAPORT, D., & WHALEY, F. Problems of interpersonal relations in research on groups. *Hum. Relat.,* 1949, *2,* 281-292.

POLANSKY, N., LIPPITT, R., & REDL, F. An investigation of behavioral contagion in groups. *Hum. Relat.,* 1950a, *3,* 319-348.

POLANSKY, N., LIPPITT, R., & REDL, F. The use of near-sociometric data in research on group treatment processes. *Sociometry,* 1950b, *13,* 39-62.

PRENTICE, W. C. H. The interruption of tasks. *Psychol. Rev.,* 1944, *51,* 329-340.

ROSENZWEIG, S. An experimental study of "repression" with special reference to need-persistive and ego-defensive reactions to frustration. *J. exp. Psychol.,* 1943, *32,* 64-74.

SCHACHTER, S. Deviation, rejection and communication. *J. abnorm. soc. Psychol.,* 1951, *46,* 190-207.

SLIOSBERG, S. Zur Dynamik des Ersatzes in Spiel- und-Ernstsituationem. *Psychol. Forsch.,* 1934, *19,* 122-181.

SPENCE, K. W. Types of constructs in psychology. *Psychol. Rev.,* 1944, *51,* 47-68.

THIBAUT, J. An experimental study of the cohesiveness of under-privileged groups. *Hum. Relat.,* 1950, *3,* 251-278.

THIBAUT, J. W., & COULES, J. The role of communication in the reduction of interpersonal hostility. *J. abnorm. soc. Psychol.,* 1952, *47,* 770-777.

TOLMAN, E. C. *Purposive behavior in animals and men.* New York: Appleton-Century, 1932.

TOLMAN, E. C. A psychological model. In T. Parsons and E. A. Shils (Eds.), *Toward a general theory of action.* Cambridge: Harvard Univ. Press, 1951. Pp. 279-359.

WHITE, R. K. The case for the Tolman-Lewin interpretation of learning. *Psychol. Rev.,* 1943, *50,* 157-186.

WIEHE, F. Unpublished monograph: Die Grenzen des Ichs. Referred to in K. Lewin, *A dynamic theory of personality.* New York: McGraw-Hill, 1935.

WRIGHT, H. F., & BARKER, R. G. Methods in psychological ecology. Kansas: Univ. of Kansas, 1950.

ZEIGARNIK, BLUMA Ueber das Behalten von erledigten und enerledigten Handlungen. *Psychol. Forsch.,* 1927, *9,* 1-85.

ZELLER, A. F. An experimental analogue of repression. I. Historical summary. *Psychol. Bull.,* 1950, *47,* 39-51.

CHAPTER 6

Role Theory

THEODORE R. SARBIN*
University of California
Berkeley

Role theory attempts to conceptualize human conduct at a relatively complex level. In a sense it is an interdisciplinary theory in that its variables are drawn from studies of culture, society, and personality. The broad conceptual units of the theory are *role,* the unit of culture; *position,* the unit of society; and *self,* the unit of personality. We define position as a system of role expectations. The discussion of this definition will attempt to show the interrelatedness of role and position.

Coordinate with role is the concept of the self. Its origins and its dimensions are described in cognitive terms. The ultimate units of the self are inferred qualities, the conceptualization of which is aided by the use of qualifying terms such as adjectives. The self-and-role-in-interaction thesis prevails throughout the chapter and we shall report empirical observations that focus on the interactional nature of conduct.

Since we hold the position that the value of a theory can be assessed only through empirical verification of hypotheses drawn from the theory, we review a number of representative studies. The results of these studies provide at least some evidence for the utility of theory.

The theory set forth in this chapter leans heavily on the concepts of George Herbert Mead (1934). It may properly be called an interaction formulation. However, it differs in certain respects from conventional interaction theories which are predicated on the notion that action occurring between persons is reciprocal. The general formula is: when *A* initiates an action to *B, B's* response to *A* serves as a stimulus for *A,* etc. The unit of action is the person. It would not be too far from the truth to say that all strictly sociological theories accept implicitly or explicitly this formula of interaction. Analysis which follows from such interaction theories would allow inferences at the first of the three levels of complexity described by Murphy (1947). At this level persons are studied as distinguishable entities without reference to internal structure.

In role theory, the person as the broad sociological unit of interaction is retained, but a somewhat finer unit, the role, is added. Thus role theory embraces reciprocal action between persons, but these actions are organized into roles. If this were the only addition, we would have no more than an extension of traditional sociological theory. A second kind of interaction has been added, however, which marks role theory as a unique social psychological formulation, namely, the interaction of role and self. Such a theory aims at problems at the third of Murphy's levels of complexity, where there is structure within the organism and structure within the environment. It is to the investigation of these structures and their interaction that role theory addresses itself. In broad perspective, contemporary role theory regards human conduct as the product of the interaction of self and role. Not dissimilar is Parsons and Shils' idea of need dispositions and role expectations (1951).

Unlike the field-theoretical school or the psychoanalytic school, whose concepts have histories more or less directly traceable to the original theorists, Lewin and Freud, role theory has developed as the result of the work of independent investigators who "discovered" the need for such synoptic concepts as self, role, and interaction. This does not imply that these conceptions have no history. At the turn of the century James M. Baldwin, G. Stanley Hall, William James, and John Dewey had begun to place the self on an empirical footing. Not long afterward Charles H. Cooley and George H. Mead further developed the notion of the self, and Mead began his analysis of the role-taking process. With a few notable exceptions, however, the concept of the self disappeared from learned psychology until it was revived by Allport (1937).

---

*I am grateful to Wallace B. Hall, Donal S. Jones and Donald G. Woodworth for bibliographic help and for critically reading parts of the manuscript. I am also indebted to Professor Paul Wallin of Stanford University for his stimulating criticisms. A special debt of gratitude is owed my wife, Genevieve, who helped in many ways, but especially in establishing the conditions which allowed me to meet the editor's deadlines.

The concept of role, until recently, was no more than an armchair abstraction and made little, if any, impression on the currents of American psychological thought. In the last decade, under the influence of the sociometrists and group dynamics workers, and as a result of interdisciplinary work with sociologists, role has been accepted as a concept with research possibilities. Interaction has always been recognized as a common-sense principle, but psychological experimentation and conceptualization were designed primarily in unidirectional stimulus-response terms.

At this writing, there is a growing number of psychologists who utilize self, role, and interaction as focal concepts in their theoretical formulations. Among these are Newcomb (1950), Cameron (1943, 1947), Sargent (1950), McClelland (1951), and Sarbin (1943). Mention should also be made of a number of sociologists whose contributions have been invaluable: Cottrell (1942), Merton (1940), Hughes (1928), Waller (1932), Parsons (1951), Parsons and Shils (1951), and K. Young (1952).

It should be noted here that most of the literature is discursive and anecdotal or abstractly theoretical. Although we shall refer at the appropriate places to some of these authors, wherever possible we shall review empirical or experimental studies in order to document our argument. Parenthetically, Cottrell's presidential address before the American Sociological Society (1950) sketches a number of crucial problems awaiting empirical investigation, most of which focus on self, role, and interaction.

The exposition of role theory requires a broad overview against which our psychological concepts may be delineated. To this end, a conceptual framework is sketched below which embraces culture, society, and personality.

With Linton (1947), we regard a culture as an organization of learned behaviors and the products of behavior which are shared and transmitted. When analyzed, these behaviors appear to be no more and no less than the ordered *actions* of persons, e.g., "mother feeds infant by forcing premasticated food into its mouth." Moreover, persons are always members of a society (defined as an aggregation of persons with common goals), and these societies are structured into *positions* or statuses or offices. The positions are collections of rights and duties designated by a single term, e.g., mother, village chief, teacher, etc. The actions of persons, then, are organized around these positions and comprise the *roles*. Role and position are conjoined. Roles are defined in terms of the actions performed by the person to validate his occupancy of the position. In sum, all societies are organized around positions and the persons who occupy these positions perform specialized actions or roles. These roles are linked with the position and not with the person who is temporarily occupying the position.

Furthermore, the person is characterized by an internal organization of qualities or dispositions, the residue of an individual's experiences as a participant in the culture. In Linton's analysis this aspect is called *personality* and emphasizes individual differences. In role theory, this internal organization of qualities (traits, attitudes, and habits) is conceptualized as the *self*. The term "personality" is reserved for a broader concept: action systems arising out of the interplay of self and role.

ROLE

Although the term role has recently come to have the beginnings of a stable core of meaning, we are a long way from univocality. In a review of over 80 sources in which the concept *role* was used, Neiman and Hughes (1951) concluded that:

> The concept *role* is at present still rather vague, nebulous, and nondefinitive. Frequently in the literature, the concept is used without any attempt on the part of the writer to define or delimit the concept, the assumption being that both writer and reader will achieve an immediate compatible consensus . . . (p. 149)

As employed by Mead (1934) and a number of other sociologists such as Park and Burgess (1924), role was regarded as the unit of socialization. This usage leads to a definition of personality or of the self as the internalization of roles. Whether the role is made up of overt or covert elements or both is difficult to say from their statements. Role is regarded by them as "a pattern of attitudes and actions" which a person takes in social situations. One objection to this formulation is that it is not conceptually pure. Attitudes are inferences made about certain kinds of covert behavior, while actions are more directly observable. A unit which includes such operationally heterogeneous conceptions makes difficult the isolation of variables for research. The lack of specificity in the above definition of role may be responsible in part for both the paucity of experimental studies and the plethora of armchair theorizing associated with it.

Whatever agreement has been attained in the use of the term *role* centers around the organized *actions* of a person coordinate with a given status or position. One of the sometimes confusing results of thus regarding roles and positions as intimately related is in their being treated as identical. They are interdependent concepts, but not identical. In the paragraphs which follow an attempt is made to translate position into psychological concepts and to indicate its relationship to role.

DEFINITION OF POSITION

What is a *position?* — or its equivalent, "a system of rights and duties"? The writer would regard a position in a social structure as a set of expectations or acquired anticipatory reactions. That is to say, the person learns (a) to expect or anticipate certain actions from other persons and (b) that others have expectations of him. The child, for example, learns to expect to be fed, bathed, fondled, etc., by an adult who, of course, is operating within an existing role organization. These expectations become organized into a concept later verbalized as "mother." The adult expects certain responses from the child. These latter expectations become organized into a concept such as "mother's darling." Another illustration is the job specification of contemporary personnel work. The job or position is described in terms of the actions *expected of* the occupant of the position (his obligations) and the actions which he may *expect from* others such as foremen, subordinates, paymaster, etc. (his rights). In other words, a position is a cognitive organization of expectations, a shorthand term for a concept embracing expected actions of persons enacting specified roles. These expectations, organized as they are around roles, may justifiably be called role expectations. Thus, a position is a cognitive organization of role expectations.

In somewhat different language but probably related to this conception is the statement of Parsons and Shils (1951). Although roles and role expectations are used interchangeably, the referent appears to be the same.

> The allocative foci of social systems are roles or role expectations. The social system is in a sense composed of a variety of roles or role expectations; each of these assures that some need of the social system will be met. The referent of a role . . . is a set of concrete orientations; the role or role expectation is an inferred entity in exactly the same fashion as is the need-disposition. When orientations are grouped (or abstracted) according to the roles or role expectations that control them, and according to the interacting groups to which they belong, we are dealing with social systems. (p. 92)

DEFINITION OF ROLE

Cottrell's concise definition is worth quoting: "A role (is) . . . an internally consistent series of conditioned responses by one member of a social situation which represents the stimulus pattern for a similarly internally consistent series of conditioned responses of the other (s) in that situation. Dealing with human behavior in terms of roles, therefore, requires that any item of behavior must always be placed in some self-other context" (Cottrell, 1942, p. 617).

A *role* is a patterned sequence of learned *actions* or deeds performed by a person in an interaction situation. The organizing of the individual actions is a product of the perceptual and cognitive behavior of person *A* upon observing person *B*. *B* performs one or a number of discrete acts which *A* observes and organizes into a concept, a role. On the basis of this conceptualization of the actions of *B*, *A* expects certain further actions from *B*. This expectation is covert, and is the equivalent of saying "locates or names the position of the other." Once having located or named the position of the other, *A* performs certain acts which have been learned as belonging to the reciprocal position; these actions are conceptualized as *A's* role. Although language imposes upon us the necessity for defining role in this rather deliberate way, the reader should bear in mind that this "locating the position of the other" may be placed somewhere on a continuum from deliberate to automatic, from witting to unwitting.

In attempting this definition of role and position, the interdependent relationship of the two concepts is indicated. A role, the organized actions of a person in a given position, may be a person's own, or it may be the actions of other persons not immediately involved in the reciprocal behavior (Cameron, 1950). The role may be that of teacher, the reciprocal role that of student, and both teacher and student may be reading about a third role which belongs to neither of them, for instance, scientist. The interest of the investigator, then, can focus upon the actions of the teacher and the role expectations of the student, or vice versa, or upon the influence of the actions of the scientist on the role expectations of either teacher or student.

ROLE EXPECTATIONS

A position in a social structure is equivalent to an organized system of role expectations. In Tolman's (1952) scheme, the latter would be variables intervening between the stimulating conditions and the performances or acts. These role expectations are cognitive and, of course, must be inferred either from phenomenological report or from the analysis of other kinds of data. Two general kinds of expectations are found: rights and obligations. Rights are role expectations in which the actor of the role anticipates certain performances from the actor of the reciprocal role; e.g., the child's right to be protected by his mother. Obligations (or duties) are role expectations in which the actor of a role anticipates certain performances *directed toward* the actor of the reciprocal role; e.g., the mother's obligation to provide protection for the child.

A person cannot enact a role for which he lacks the necessary role expectations. These must be acquired through experience. In the following pages we shall consider the ways in which persons acquire the prior experience which makes possible the enactment of roles.

THE ACQUISITION OF ROLES

The actions which are patterned into roles are learned through two broadly defined processes, *intentional instruction* and *incidental learning*. It should be explicitly stated that both forms of acquiring roles may operate conjointly. Any set of role expectations is probably an admixture of expectations acquired through intentional instruction and incidental learning. In intentional instruction the culture carriers in any society attempt to teach prescribed acts. Intrinsic to these methods is the belief in some form of reinforcement theory of learning. Some cultures stress reward for the commission of a desired act, others stress punishment for the commission of an undesired act. In either case there is an intentional program designed to teach the child certain role behaviors.

Early in life the child is taught acts which are differentiated, for example, according to sex. Little boys are rewarded or punished for certain kinds of actions, little girls are rewarded or punished for other kinds of actions. Thus begins the acquisition of actions for the later performance of one's generalized sex role.

Acts are learned as patterned behaviors, not as isolated muscle responses. Further, these acts are taught through the agency of other persons; thus, stimulus events always include social objects who in turn respond to the child's acquired actions. This is the paradigm of interaction for which the role is the unit. The child learns not only the motoric actions appropriate to the given situation, he also learns that the other person has expectations into which his overt acts must fit. Concurrently, in the sequential interactions, he comes to expect certain responses from the other. Thus is born his own system of expectations.

In incidental learning, the child adopts the ways of others in his environment. Cameron and Magaret (1951) refer to this as "adoption of the prevailing pattern." DuBois (1941) refers to "absorptive" disciplines in child rearing.

Several forms of acquiring role behavior through incidental means may be identified. Among them, play is probably the most widely observed. Mead (1934) was among the first to recognize the importance of play in the adoptive learning of social roles. At least two resultants follow from the play acting of children: (a) the acquisition of roles (truncated, of course, because of maturational limitations), and (b) the acquisition of skill in shifting roles. In play, the child can shift from role to role without observing the formal logic of the adult. This movement from role to role leads to oscillatory shifting of sets, thus enabling the child to take both his own role and the role of the other. As will be shown later, the absence of a number of different standard roles, as well as the absence of skill in taking-the-role-of-the-other, retards socialization and leads to invalid role enactments.

The imaginative processes are central in play acting. They are likewise central in covert processes, such as fantasy. The silent rehearsal of roles appropriate to real or imagined positions, and of roles appropriate to the position of the other, provides a large reservoir of experience. In all cultures, a good part of the content of imaginative behavior is institutionalized in the form of folk tales, myths, and other story forms. The imaginative process is central likewise in that form of acquiring roles which has been variously named identification, introjection, empathy, and taking-the-role-of the-other. Dependent on the ability of the person (child or adult) to engage in *as if* processes, identification provides numerous avenues for acquiring roles. The number and kinds of persons with whom one may identify, of course, is limited by the number and kinds of persons in the environment and by cultural practices. If a child's social environment is made up principally of the mother, there are fewer opportunities for identification than if the environment contains many persons. The "other" need not be in the visual field of the acting person. In fact, he may be a cultural hero, contact with whom has been established vicariously through storytelling. The importance of the imaginative processes in role taking has been emphasized by Mead (1934), Moreno (1946), Cameron (1947), and Sarbin (1950).

Superficially similar to identification, and probably antecedent in the development of the child, is the deliberate copying of the actions of others in the absence of intentional instruction. Although imitation as an explanatory device has been repudiated, the observations that children deliberately copy the behavior of other children and adults cannot be gainsaid. What is copied, of course, depends on the child's maturational level, availability of models, and favorable motivation (Miller and Dollard, 1941). Unlike identification, the model has to be present in the imitator's behavioral field.

Controlled experimentation in the learning of roles has not yet begun. Anecdotal material is available which provides evidence for the plausibility of the hypothesis that adult role behavior depends on the prior acquisition of role expectations. Especially cogent are the reports of situations in which the absence of the opportunities for learning to enact roles appropriate to defined positions leads to deviant conduct (Cameron, 1947; Gough, 1948).

ROLE EXPECTATIONS AS ACTIONS AND AS QUALITIES

Both kinds of role expectations, rights and obligations, may be analyzed in terms of *actions* and *qualities*. Analysis of the role expectations of mother, for instance, may reveal the potential action: mother provides food when child cries. In addition, certain *qualities* of the person enacting the role are expected: warmth, gentleness, friendliness, etc. Both the expectations of actions and the expectation of qualities are residua of prior experience. This construction is similar to Parsons and Shils' notion of performance and qualities.

> It is possible to orient objects either (1) in terms of characteristics they possess regardless of their performances, or (2) in terms of characteristics they possess by virtue of their performances . . . A social object is a complex of qualities when the actor overlooks . . . performances and focuses on "attributes" as such . . . A social object is a complex of performances when the actor . . . *focuses* on the processes of action and their *outcomes* rather than its qualities or attributes . . (1951, p. 65)

Role expectations as actions. Role expectations as anticipated actions or performances can be studied by means of questionnaires. A number of questionnaires appropriate to such study have been devised. One of these, a role-conception inventory, has been developed by Motz to determine (1) the expectations of actions of self in the marital relationship, and (2) the expectations of actions that other occupants of the same position *should* perform. Preliminary interviews with married students indicated the presence of two types of marital roles for each sex: traditional and companionship roles. These types were conceptualized according to six areas of *actions:* housework, employment, financial support, child care, community participation, and schooling. Statements in the inventory were devised to get at these action expectations in the traditional role and in the companionship role. The items were so written that scores could be obtained for personal expectations ("I should do thus and so") and public expectations ("Wives should do thus and so"). The variability in expectations among the members of the group, of course, reflects the degree of clarity of role expectations.

Cottrell (1942) has listed a number of hypotheses regarding the relationship of role expectations to adjustment. One of these hypotheses is concerned with the specificity and definiteness of role expectations. "The degree of adjustment to roles which a society assigns to its age-sex categories varies directly with the clarity with which such roles are defined." As the result of a questionnaire study of undergraduates, Rose (1951) concluded that the adult role is unclear to the average college girl. Role expectations about children, household chores, servants, civic work, employment, etc., showed lack of definitions. That such a state of affairs may lead to ineffectual role enactments is a safe guess. A person must move cautiously and uncertainly when role expectations of others are partly known or entirely unknown (or unknowable). Role-role and self-role conflicts are likely to follow from ambiguous role expectations. The persisting need for solution of such conflicts may lead to socially invalid role enactments.

Lack of specificity and definiteness in role expectations may be related to discontinuities in female expectations. Benedict's (1938) often-quoted observations about the effects of discontinuity in cultural conditioning are apposite here. The analyses by Komarovsky (1946), Wallin (1950), Nottingham (1947), and Hacker (1951) on the effects of historical changes on role expectations emphasize the actual and potential problems of women in contemporary American society.

Emphasis on the lack of specificity and definiteness in women's role expectations should not lead the reader to assume that this condition does not apply elsewhere. Persons who are mobile from class to class or culture to culture may likewise be plotted on this dimension. The American adolescent's occupational role expectations are likewise characterized by indefiniteness and diffuseness (Carter, 1944).

Role expectations as the independent variable in conflict analysis has been used by several investigators. Role conflicts occur when a person occupies two or more positions simultaneously and when the role expectations of one are incompatible with the role expectations of the other. Ort (1952) used the concept of role expectations in an empirical investigation of the relationship of social class to role conflicts. The subjects were 100 adolescent boys, and the data were determined simply by counting the number of unfulfilled role expectations. An example of an unfulfilled role expectation would be where a youth answers "Yes" to the question "Should a boy be home by a certain time at night?" and "No" to the question "Are you home by a certain time at night?"

Social class indices were derived from Warner's scheme. The correlation between role conflict and class was positive and high; as predicted, the middle-class adolescents expressed more role conflicts than the lower-class boys.

This research is cited only incidentally to point out the influence of class position on role conflict. The significant point is that it demonstrates how a cognitive variable, action expectations, is subject to empirical research.

Less frequently used than the questionnaire or the interview, but perhaps possessing more ecological validity (Brunswik, 1947), are behaviors observed in enacting contrived roles. Inferences are drawn from the observation of a person actually engaged in the enactment of a role. Of course, the investigator must make a correction (implicitly) for the fact that the subject, in addition to responding to the task set for him, is also enacting the role of "subject-in-an-experiment."

Brown (1952) reported an experiment in which this role-playing method was used. He was interested in the effects and interactions in role-taking situations in terms of the following characteristics: sex of person enacting the role of another person, sex of the "other" (the reference person), the sex orientation of the cultural content of the situation, and incidental learning. His main hypothesis was that "there exists a partial disjunction between the cultural worlds of males and females in American society which will be exhibited in the partial failure of role taking" (Brown, 1952, p. 568). In other words, the results of learning role expectations of males and females are such that role enactment of an opposite-sex peer will be inadequate or distorted. In general, Brown concluded that (a) males are slightly better than females in role-taking skill and that (b) role taking across the sex line is more difficult than role taking within sex categories.

Role expectations as qualities. Role expectations as qualities or attributes rather than as actions or performances may also be studied empirically. The significance of inferred qualities in building up role expectations is derived from a study by Sullivan (1950). He investigated conceptual behavior with the aid of a set of 40 pictures of men. The subjects were asked to sort or categorize the persons. The principles upon which the sortings were made were analyzed. Among the findings were sex differences in the way social concepts are formed. Women differed from men in that they were more inferential, more dichotomous and specific in the manner in which they assign qualitative role expectations to others. For example, women used traits (e.g., moody) and expressions (e.g., quizzical) as concepts; men used more descriptive categories (e.g., directions, clothing).

Sullivan's results suggested a further hypothesis: that women would use inferential terms depicting "underlying" traits more frequently than men in their role expectations of "live" subjects. Conversely, men would be expected to use terms more continuous with the raw data: nouns describing positions, or adjectives describing external characteristics of the person. In a study conducted by the author as a class exercise, 131 upperclassmen served as subjects. They were told that a version of the word-association experiment would be performed. Instead of presenting a word as a stimulus, however, a person would be the stimulus. The subjects were to write the first three words that "came to mind" upon the presentation of a live social object for ten seconds. This was repeated with two other persons as stimulus objects. The social objects merely rose from their seats — being required to say nothing. The descriptive words used by the respondents were then grouped into two categories: those representing several degrees of inference, such as shy, aggressive, warm, hostile, logical; and those representing descriptive categories, such as student, tall girl, grey (color of clothing), etc. The hypothesis derived from Sullivan's study was verified ($P = .02$), men favoring the use of the latter class of words, women, the former. Thus, qualitative features of role expectations may be approached experimentally (Sarbin, 1952b).

Naboisek (1953) made a start at a more ecological method of isolating qualitative characteristics of role expectations by predicting, in a limited way, interpersonal behavior to be expected from members of small therapy groups. The subjects filled out personality inventories and adjective checklists made up of 96 interpersonal adjectives. Naboisek then classified the patients according to what he considered the prevailing qualitative expectations. Four classes of role expectations were used: friendly-dominant, hostile-dominant, friendly-submissive, and hostile-submissive. After six sessions of group therapy (a reasonable time in which to become acquainted) each subject checked interpersonal adjective checklists for every other member in his group. The assessments were made about other persons enacting the patient role each in his idiosyncratic way. The agreement among their assessments verified the predictions made by the experimenter prior to therapy. For example, on the basis of pretherapy self-descriptions a person was classified as having role expectations of others as *"friendly-dominant."* After six weeks' acquaintance a summing of his checklists about the role enactments of others revealed that he actually checked adjectives about the others that were classifiable as friendly-dominant. In this setting, the roles are described qualitatively in adjectival terms, the substantive "patient" applying to all members of the group.

Developmental studies in role expectations. Genetic accounts of role expectations (either as anticipated actions or as qualities) are virtually nonexistent. Hartley and associates (1948) report a set of experiments on role expectations. By the use of pictorial materials and interrogation procedures, data were gathered to answer questions about multiple group membership. Children from age 3.5 to 10.5 were the subjects. Role expectations of father and worker were studied. Constancy of the parental role appears at an earlier age than constancy of the worker role. Whether the earlier constancy of the parental role is due to affective factors (child's *need* for a parent) or cognitive factors (greater familiarity) is a problem for future investigation.

ROLE PERCEPTION

If we accept the premise that perception is an intraorganismic response of the organism to stimulus objects and events, then the extension to the social act is obvious. The perception of roles is an organized response of a person to stimuli in a social context. This is no place to review the literature on *what* is organized. Most psychologists agree that what is organized is a contemporaneous event and the residua of the organism's prior experience. This may be conceptualized as trace and process, concept and percept, hypothesis and information, creed and deed, expectations and acts, etc. In role-taking theory, acts of persons are seen as organized against a cognitive background of role expectations. Role perception may be thought of as a sequence of behaviors in which the perceptual response is the first part of a social act: the (usually) silent naming or locating the position of the other (from observed actions or inferred qualities), which serves to locate the position of the self. The second part of the social act is the motoric response, the role enactment, in which the actor performs actions appropriate to his location of the positions of self and other. In real life situations, the sequence of events is seldom as deliberate as this analytic description would suggest. The time elapsing between the perceptual response and the motoric components of the social act may be infinitesimal, so that observers may overlook the cognitive phase. In infants, of course, the two phases are not at all discontinuous. Perception of an object or event and motor discharge are continuous processes. In young children, the motoric response which follows the presentation of a stimulus object is mediated by some form of gestural activity, such as pointing or naming. With the muting of speech, the perceptual response becomes covert.

From this presentation of role perception, it becomes obvious that the veridicality of one's role perceptions and functional adaptation to the social world are highly related. If a person's locating of the position of the other is invalid, then his location of the position of self is likely to be invalid. His role enactment, then, will be inappropriate and nonadaptive.

Research in this area has been neglected. Perception has, for the most part, been studied in "impoverished" settings, where there is little ecological generality (Brunswik, 1947). Moreno (1946) has suggested procedures which utilize social acts in the psychodramatic experiment. However, these procedures have been used more for demonstration and clinical purposes than for the determination of principles of role perception.

The building blocks of role perception, of course, are the overt acts of others. These acts may be defined as everything the other does. This includes his gross skeletal movements, his verbal behavior, what he wears, his facial expression, his posture, his gait, his accent and intonation, his adornments, visible emblems and tattoo marks, etc. The wearing of a uniform and a silver badge, for example, is part of the role of policemen. The perception of these objects on a person leads to predictable role behavior on the part of the perceiver (providing the system of role expectations of the perceiver is known). Perceiving the position of the other, then, becomes a matter of attending to and organizing cues.

Two recent experiments illustrate a method for studying role perception as well as providing data for testing certain hypotheses. In the first experiment (Sarbin and Williams, 1953) we addressed ourselves to the problem of role perception on the basis of certain reduced cues, namely, single verbal sentences. We recorded 38 sentences spoken by different persons under various kinds of role instructions. For example, one sentence was shouted by a middle-aged man: "All right, boys, get in there and fight!" Another example is a whining little girl who cries: "Somebody broke my dollie." A third example is a man saying: "I used to sit before the fire and read poetry for hours."

These recordings have been played before various groups over the past three years. The first of our groups received instructions (a) to indicate age, sex, and social or occupational role of the speaker and the person spoken to, and (b) to specify the action occurring. On the basis of 96 protocols, a multiple-choice answer sheet was developed which provided (a) three age-range choices, (b) male and female choices, (c) four role choices, (d) four "alter" choices, and (e) four action choices. Except for (b), which was eliminated as a variable early in our experiments because of the restricted range of choices, the frequency distribution for each variable followed a *J*-curve.

The first finding of importance is that different middle-class groups perceive these restricted role cues in much the same way. It was possible therefore to establish a modal response for the four response categories for each of the sentences. Although the objective validity of the response was practically zero, the consensus was high. Actually, the sentences were spoken by student and adult actors. Few of the respondents perceived the speaker as "actor." In the first illustrative sentence cited above, over 90 percent of our respondents checked "football coach."

One of the hypotheses confirmed by the experiment was that "good" role perceivers (those who respond in the modal way) would more likely have self-characteristics in common, while "poor" role perceivers (those who respond in deviant ways) would have fewer self-characteristics in common (more heterogeneity), and would be different from the "good" role perceivers. Self-characteristics were inferred from adjective checklists filled out by the subjects.

No claim is made for the objective accuracy of the self-descriptions. They are limited phenomenological reports, qualitative expressions of the self. The adjectives characteristic of "poor" role perceivers (male) are similar to the cluster checked by diagnosed neurotics (Sarbin and Rosenberg, 1952); the adjectives checked by the female "poor" role perceivers are often utilized in self-descriptions by noninsightful hysterics. The adjectives checked by the "good" role perceivers indicate awareness of tension-binding qualities (e.g., self-denying, tense, modest, thoughtful, anxious, civilized, deliberate, patient, reasonable, etc.). Obviously, this conclusion is only suggestive and calls for further testing.

A third finding emerged from the test of the hypothesis that males and females will differ in role perception. For women a significant positive correlation is found between frequency of perceiving the modal *age* response and frequency of perceiving the modal social *role* response. Age and social role thus appear to be intimately related for women. For men a significant positive correlation is found between the frequency of perceiving the modal social role and the modal social *situation*. Thus, for men, social role and social situation appear to be intimately related.

A second experiment (Sarbin and Hardyck, 1952) followed the same design as the previous one. This study was designed to investigate the influence of postural behavior (as actions of persons) on role perception. The idea of postural cues, of course, goes back to Lipps' doctrine of kinesthetic inference. We were able to eliminate cues other than postural by the simple expedient of representing postures with line drawings of "stick figures." We prepared 42 such figures and presented them by means of lantern slides to various undergraduate classes. The respondents were to write the action, state, feeling, or other quality that was suggested by the stick figure. From the responses of the first group, we constructed a multiple-choice type test for administration to other experimental groups. As in the experiment on auditory cues, we found that for most of the stimulus presentations, the typical curve of the distributions of the responses was a *J*-curve. For our sample of undergraduates, then, we can conclude that certain postures are perceived as associated with certain acts or with certain qualities of a person. Figure 1 shows a few samples and response choices.

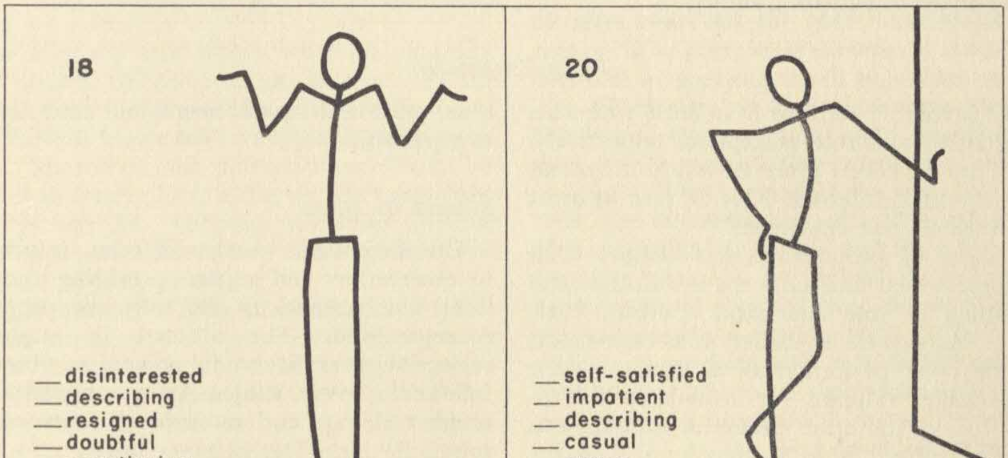


FIGURE 1.

We were interested in testing the hypothesis that persons who do not conform to the mode in the perception of postural acts differ on other criteria. For one sample the subjects were asked to state briefly the method they had used in forming their judgments, i.e., whether they had imagined themselves in the particular postures for each figure (empathized), or whether they had formed their judgments solely on the basis of the checklist categories.

An initial scoring using simply the number of modal responses for each subject indicated a significant relationship between modal response score and self-report on empathy-nonempathy. The subjects who had checked the greatest number of modal responses reported the use of the empathetic method in a significantly greater proportion, the biserial *r* being .41, $P = .01$.

For another group, scores on the California Psychological Inventory (Gough) were ranked and compared with a ranked frequency count of the number of modal responses. The relationship is quite high, the Spearman rank difference correlation being .81. In view of the nature of the California Psychological Inventory, social adjustment would seem to be definitely linked to empathic ability.

A third procedure was to administer the stimulus series individually to persons whose actions had already demonstrated their lack of conformance in the perception of roles, namely, hospitalized schizophrenics. They checked the nonmodal responses as frequently as the modal responses.

Most significant in this procedure is the fact that role perception is subject to empirical research. Three different samples show relationships which are predicted from the theory elaborated below.

ROLE ENACTMENT

The preceding sections have dealt with role expectations and role perception, respectively. These are the covert processes which, of course, must always be inferred. Now we turn to overt performances, role enactment.

A series of propositions has already been advanced which facilitates the description and exposition of role enactment (Sarbin, 1943, 1951). As expressed earlier, role enactments are the overt performances of persons; these performances validate (or invalidate) the expectations of the other person or persons in a social situation. (Our interest here is in the motoric aspect rather than in the cognitive aspects of role theory.) Coutu (1951) has recently inveighed against the confusion of the motoric and the cognitive aspects. What we call role enactment is equivalent to Mead's and Coutu's "role playing." "Role taking" or taking-the-role-of-the-other means adopting the perspective or attitude of the other. This is regarded as a covert process, although there are notable exceptions, such as the overt empathizing with a halfback making an end run, smiling with the hero when he sees his sweetheart after a misunderstanding, etc. In these situations, the cognitive behavior "spills over" into action. In our organization of the theory, we treat empathy as a separate concept, as primarily covert, although we recognize its close relationship to role enactment.

Role enactment [equivalent to Newcomb's "role behavior" (1950)] includes, among other segments of behavior, gross skeletal movements, the performance of verbal and motoric gestures, posture and gait, styles of speech and accent, the wearing of certain forms of dress and costume, the use of material objects, the wearing of emblems or ornamentation, including tattoos, etc. In short, role enactment embraces what may be called the mechanics of the role-taking process. These mechanics can be summarized by means of the following concepts: number of roles, organismic involvement, and accessibility or reportability.

NUMBER OF ROLES

The dimension, number of roles, is subject to observation and report, providing the actions which make up the roles are properly conceptualized. The ultimate in attaining ecological generality would occur if an observer followed a given subject for a representative number of days and recorded all instances of role behavior. The observer might also rate the intensity with which he enacted the various roles. This type of observational approach is not new to anthropologists (see, for example, Ford, 1938).

A less difficult approach for getting at the number of roles enacted by an individual is the use of small-group settings. The discussion group has been the focus of a number of research attempts in recent years and lends itself to the laboratory study of role behavior. Benne and Sheats (1948) have recognized and named a number of positions (systems of role expectations) which characteristically arise in the discussion group. Some examples are: the opinion-giver, the orienter, the encourager, the hostile critic, the aggressor, etc. It should be pointed out that such roles do not have the institutional support of the more traditional roles, such as mother, doctor, shortstop, etc. Nevertheless, they do emerge in the interaction process and can be described with the same formal characteristics as the traditional roles. Among the roles which emerge in the small group are various "leader" roles, such as initiator, coordinator, director, and so on. The dimension "number-of-roles" was applied to such leader roles in a recently completed study by Sarah Mazelis (1953). She addressed herself to the question: Is the productivity of a discussion work group related to the diffusion of leader roles among the members of the group? The leader roles were defined and categorized by actions. The observers recorded and counted the number of leader roles enacted by each person in each of four independent work groups. The hypothesis studied was: productivity will be greater among groups where the various leader roles are enacted by most of the group members; productivity will be less where the leader roles are enacted by a small proportion of the group members. The work products were rated on several criteria by experts in the field.

In a group task which required planning, the quality of the product, as rated by the judges, was directly related to the diffusion of the various leader roles among the group members. Spearman rank difference correlation between an index of role diffusion and production ratings was .75. Although the small number of groups makes this finding only suggestive, the method appears promising.

This number-of-roles dimension is implicit in the application of the role-taking approach to psychopathology. It is a widely accepted postulate that the more roles in a person's behavior repertory, the "better" his social adjustment — other things being equal. Cameron and Magaret (1951) have argued most convincingly that the absence of role-taking skills is influential in the development of paranoid disorders. Gough (1948) has also implicitly used number-of-roles as a dimension in his role-taking theory of psychopathy. Following Mead, he describes how the residuals of role enactments ("me's") are integrated into a self-conception, and how the ability to be self-critical is a function of the number of "me's" or the internalization of a "generalized other." Further, ". . . deficiency in role playing means the incapacity to look upon one's self as an object (Mead) or to identify with another's point of view . . ." (Gough, 1948, p. 363). Gough and Peterson (1952) have already utilized this conception in an empirical study of psychopathology.

ORGANISMIC DIMENSION

Any role may be enacted with different degrees of organismic involvement. This is essentially an intensity dimension, the intensity of the enactment being manifest in the number of organic systems involved. At the low end of the dimension would be the kind of interaction which occurs with little affect and with little effort. The role of the customer in today's supermarket involves only minimal participation — the saying of a few words, plus a few movements involved in the exchange of money. Contrast this role with the intensity involved in enacting the role of the mother of a sick child.

In daily life, of course, most roles are enacted with minimal organismic involvement. Behavior would be inefficient indeed if all roles were enacted with maximal intensity. Not only are cultures organized so that the number of maximally intense roles is few, but there are autonomic safeguards which limit the time span of roles that call for great mobilization of energy. The organismic dimension can be understood as applying to all organized actions. The illustrations we have chosen are reference behaviors against which more conventional roles may be compared for intensity (Table 1).

Level I. The low end of the scale is labeled "role and self differentiated." The self, a cognitive structure (*v. infra*), and the role, also a cognitive structure, are minimally in contact. An example of a casual role, already given, is the customer in a supermarket. In fact, it is possible to perform this role at the motoric level while at the same time vicariously enacting other roles unrelated to the public role.

Level II. At the next level of organismic involvement, the reference is that of the dramatic actor who performs the motions necessary for the portrayal of the role assigned to him. In mechanical acting, the actor does not become involved; the self is relatively autonomous from the role. He must maintain a certain degree of consistency in the various response systems, which calls for more effort and precludes the degree of autonomy of the casual role. This degree of involvement is seen in many everyday interactions. For example, the employee who "puts on a good front" to impress the boss.

Level III. At the third level, "heated" acting is used as the illustrative reference behavior. Archer (1889) in his classical study of the psychology of acting, was able to discern this kind of acting from the more automatic and mechanical mentioned before. Commonly called "living the role," the actor behaves as if he *is* the character in the drama. Of course, the successful actor maintains some contact with his role as actor in order to change his tempo, amplitude, intensity, etc., as conditions warrant. His involvement includes some affective as well as motoric components. In order to portray anger, for example, the actor may work up a rage by violently shaking a ladder in the wings before appearing on the stage.

Level IV. The role of the hypnotic subject

TABLE 1

Role and self differentiated	Role and self undifferentiated
Minimal involvement	Maximal involvement
Few organic systems	Entire organism
Little effort	Much effort

I. casual roles

II. dramatic role, mechanical acting

III. dramatic role, heated acting

IV. role of hypnotic subject

V. hysterical fugue, role of the amnestic

VI. roles involving ecstatic states; mystical experiences, possession, religious conversions, sexual climax

VII. role of the moribund person; object of sorcery and witchcraft (sometimes irreversible)

serves as a reference role for a moderate degree of organismic involvement. The classical behaviors of the hypnotic subject, the catalepsies, the compulsive post-hypnotic actions, sensory and motoric changes, and so on, illustrate more cogently the operation of the *as if* mechanism. In enacting the role of the hypnotized subject as perceived against a background of generalized and specific expectations, a person demonstrates that more of the organism is responding than in play acting. He behaves *as if* he is blind, or deaf, or fearful, or analgesic, or whatever the specific instructions call for.

The hypnotic situation provides a suitable laboratory model for the demonstration of organismic involvement. An experiment conducted by Lewis and Sarbin (1943) demonstrates this involvement. These investigators addressed themselves to the question: Can *as if* behavior, as observed in the role taking of hypnosis, influence a basic physiological process, namely, the hunger cycle? Ten volunteer subjects, demonstrating different degrees of hypnotizability, participated in the experiment. The apparatus was essentially the same as used by Carlson (1916) in his studies of gastric hunger contractions. Each subject appeared for the experimental period each morning without his breakfast. He was hooked up to the recording apparatus. He was hypnotized and given no further instructions. One full gastric hunger cycle was recorded. During the middle of the second cycle, he was given a fictitious meal. The results indicated that for subjects rated high on the role-taking aptitude (see below), i.e., the ability to participate in *as if* behavior, the gastric hunger contractions stopped. That is, taking the role of the eater was effective in reducing or eliminating the gastric contractions even though no food was ingested. For subjects who were not hypnotizable, who were limited in their ability to participate in this form of *as if* behavior, no effect was produced on the gastric cycle. Thus, a range of the dimension of organismic involvement was demonstrated.

Level V. The behaviors subsumed under the rubric hysteria may be used as a reference band on the organismic continuum. Responding *as if* one is afflicted with some organic dysfunction is the main characteristic of the hysteric. The range covered by hysteria overlaps considerably the range covered by hypnotic role taking. The organismic components in hysteria, however, are less self-limiting and more prolonged than in hypnosis. Descriptive cases may be found in any textbook of abnormal psychology.

Common examples are hysterical seizures, hysterical anorexia, hysterical paralysis, and hysterical anesthesia. Both major varieties of hysteria, inactivation and autonomy (Cameron, 1947), show how the role of the invalid may involve the entire organism. In the same range of intensity of involvement is the couvade. In this custom the husband lies-in; the essence of the custom is the husband's taking the role of the wife (Crawley, 1902).

Level VI. Ecstasy, a condition usually involving suspension of voluntary action, illustrates organismic involvement to a degree which is not ordinarily observed in day-to-day social interaction. Obviously, such states cannot be prolonged over time without damage to the functioning of the body. Ethnologists' accounts of ecstatic trance experiences, possession, religious revivals, conversion experiences, and mystical unions show many of the characteristics of hysteria. But in addition to the greater involvement of the skeletal musculature, there is a greater involvement of the organs that are served by the autonomic nervous system.

In our own wider culture, such roles are taken by participants in revival meetings, by adolescent bobby-soxers who swoon upon hearing "the voice," by participants in sexual congress, by marathon dancers, etc. What is common to these events is the intensity of involvement, the apparent relationship of this intensity to the activities of the sympathetico-adrenal system, and the automatic equilibratory controls. All these activities are terminated through institutionalized rituals, fatigue and exhaustion, and/or autonomic regulation.

In *rites de passage* the role of the celebrant is characterized by high organismic involvement. The manifest purpose of the intense role behaviors of the ritual is to signify the change from one position to another in the society; the effect of the intensity of role enactment is to modify the participant's self-concept so that the new role, e.g., adult, may not be incongruent with the self. If, for example, strength is an expected property of the adult person, and the rites of passage from adolescent to adult includes passing a test of strength, then the successful completion of the test allows the person to add the adjective "strong" to his self-description. Thus he is better equipped to occupy the position of adult, not only because others know he has strength, but because he conceptualizes the self as strong.

Level VII. The ultimate limit of the intensity dimension is an extension of the previous range. The effects of the sympathetico-adrenal system which is excessively activated in such role-taking sequences as described above may become irreversible under certain stimulus conditions. The end result of such a performance is death.

Well-authenticated cases of *Voodoo* death allow the employment of the descriptive phrase, taking-the-role-of-a-moribund person. Devereux (1939) has reported of the Mohave Indians that a person who learns that he has violated the incest taboo takes the role of the person who is about to die. Rivers (1924) reports of the Papuans and Melanesians: "Men who have offended one whom they believe to have magical powers sicken and die, as the direct result of their belief; and if the process has not gone too far they will recover if they can be convinced that the spell has been removed" (Rivers, 1924, p. 50).

The case of a man dominated by a death wish is described by Alexander (1943). Autopsy findings revealed no organic pathology. Apparently he took the role of the dying person so intensely that the equilibratory processes failed to work.

The eminent physiologist, Cannon (1942), has provided a description of the physiological and social events associated with taking the role of the doomed person. He uses the description of bone pointing among the Australian aborigines provided by Basedow (1925). This is a form of black magic in which the sorcerer points a sharpened bone at the intended victim.

Cannon explains the physiological events arising from the adoption of the victim's role. The fear response acts upon the soma in many ways. The sympathetico-adrenal system comes into action, exercising control over the viscera and blood vessels. The various internal changes render the organism ready for physical action (struggle or flight). If this state of extreme perturbation continues *without motoric discharge,* death may follow. But why do the equilibratory or internal constancy mechanisms fail to re-establish homeostasis, as in ecstatic experiences? One answer to this question is that the perturbation in the organism is maintained by the enactment of supporting roles by other members of the community (Warner, 1937).

These levels of organismic involvement are equivalent to, if not identical with, the concept of self-involvement. At one end of the continuum, we see relatively automatic responses, stereotyped habits, autonomous functioning of self and role. At the other end, we see the unitary functioning of self and role.

Dimension of accessibility. Although it would be simpler to use the descriptive term "consciousness of role enactment," the writer prefers the term "accessibility," or reportability. The conscious-unconscious polarity carries with it connotations derived from psychoanalytic theory which assign motivational properties to so-called unconscious contents. Cameron (1947) has used the expression "not accessible to one's self-reactions," which carries the meaning that a person cannot make valid observations of his own role enactments.

This dimension applies to events in which the actor of the role reports his *actions.* One can immediately think of roles which are enacted automatically and where the person cannot recall his actions. The illustration given before about the person enacting the role of the customer in the supermarket can be applied here. He may perform all the actions necessary toward the selection and purchase of a wide variety of articles but be unable to report his actions. Another example is that of the hysterical amnesia victim. In enacting this role the patient does not recall events prior to the amnestic period. In Archer's (1889) study of acting, such a dimension was isolated. Some actors reported being relatively "unconscious" of self in certain roles, while in others they were extremely observant of their own role enactments.

The entire psychology of observation and report is as applicable to one's own role behavior as it is to the role-behavior of others. Research hypotheses in this area could profit from the use of the observations which the depth psychologists have conceptualized as repression and suppression.

In what is commonly called self-consciousness, the actor attends almost exclusively to his own actions rather than to the events occurring in the broader social field. The self-conscious person can report such events as the quivering of his voice, the shaking of his knees, the drying of his throat, etc. It is obvious that being greatly involved in the organismic features of the role enactment limits the possibility of observing the self as a social object. In general, one would expect to find a strong inverse relationship between the organismic dimension and the accessibility dimension. The two dimensions, however, should be characterized independently. Events that can be placed on the organismic dimension are available to the observation of others without the necessity of phenomenological report, while the accessibility dimension can be applied only to the actor's report of his own role behavior.

This dimension has not been studied empirically except in clinical settings and in hypnotic experiments. The value of this dimension of role enactment would lie in its application to the learning of social roles. Self-criticism of one's own actions could follow only if these actions were accessible or reportable. The therapeutic dictum, "making conscious what is unconscious" to effect behavioral changes, refers to the same events as making accessible to one's self-reactions inaccessible aspects of one's role enactments in order to modify or eliminate these actions. Bowman (1949) has utilized this notion in his description of the development of insight by virtue of practice in role enactment using the methods developed by Moreno (1934).

The as if dimension. It is by now a truism that the learning of social roles — in fact the learning of any concept — is associated with the ability to treat an object or event *as if* it is something else. One has to adopt the *as if* set in order to group apparently disparate objects or events into a common concept. A full philosophical and logical treatment of the *as if* principle is given in Vaihinger (1924).

In the *if statement,* or the conditional clause, something unreal, untrue, problematical, or fictive is stated. With this assumptive behavior is associated the *maintenance* of the assumption in spite of its fictive character. The maintenance of the assumption is possible only if the organism has developed verbal symbols and the capacity to bind time and tension. The imaginal processes, fantasy, identification, etc., are now seen to depend upon the capacity to engage in *as if* behavior. Between the *as* and the *if,* grammatically speaking, lies a whole sentence. The person must "hold the thought" or maintain the set established by the conditional fictive clause, while at the same time considering the independent clause. As stated earlier, this conceptual process cannot be performed if the person lacks the ability to bind time and tension, a function of the development of the self. Since roles are organized concepts, taking-the-role-of-the-other, then, is not possible without the *as if* ability.

This *as if* behavior can be treated as an aptitude or skill in the same way that we treat academic aptitude or skill. Operative in apparently diverse kinds of conduct, the *as if* formulation has been applied by the author (1950) to repression as an ego defense, impunitiveness as a response to frustration, and hypnotizability as a personality trait. These three variables (the "triadic hypothesis") have been found to be intercorrelated (Rosenzweig and Sarason, 1942). In repression the subject behaves *as if* an event threatening to the maintenance of the self-concept has not occurred; in impunitiveness the person acts *as if* the frustrating event were no longer frustrating; in hypnosis the person behaves *as if* certain qualities and events are different from what they actually are.

The *as if* skill applies to drama, hypnosis, fantasy, play — in fact, to all imaginative behavior. To test whether *as if* behavior has any appreciable effect on the performances of a subject, we can observe the overt action of persons under instructions to behave *as if* something is or is not occurring. Some experiments along this line have been reported. Jacobson (1938) and Schultz (1932) demonstrated the influence of the subject's imagining certain events upon bodily functions. Common experience illustrates how implicit imaginal processes are reflected in overt, muscular movements during the act of imagining. A person can feel his ears becoming red and his face becoming flushed when he imagines a former embarrassment. Similarly, he may experience nausea in imagining something extremely unpleasant or disgusting.

In an experiment on suggestibility Arnold (1946) relates hypnotic and suggestion effects to vividness of imagery. Her experiment allowed the inference that the more vivid the imaginative process, the more pronounced the overt movement associated with the imagined act. In a role-theoretical analysis of hypnotic phenomena, Sarbin (1950) has presented some evidence to support the hypothesis that enacting the role of the hypnotized person is a function of the veridicality of role perception, congruence of the role with the self, and role-taking aptitude. The last-named item is defined as success in the use of *as if* behavior.

As in the case of the organismic dimension and the accessibility dimension, we have no readily available measuring instruments. Of necessity we must approach this dimension indirectly. Imaginal vividness can be studied only with the aid of phenomenological report, and we are never too sure about contamination with linguistic ability in reporting details of the covert response.

As if skill has been assessed indirectly through testing the ability of respondents to predict the test-taking behavior of the other. The procedure is as follows: Smith fills out a personality test. This may be a questionnaire, trait list, attitude scale, role inventory, or other paper-and-pencil device. Smith, or Smith's products, such as a handwriting sample, is observed by Jones. Then Jones, usually on the basis of a limited sample of Smith's behavior, fills out the test as he thinks Smith would, *as if* he were Smith. This test may then be extended to Smith's predictions of Jones' predictions of Smith. In the first application of the test, Smith predicts the other's performance; he takes the role of the other *vis-à-vis* a test. In the second application, Smith predicts how he (Smith) is perceived by Jones; he takes the role of the other, *vis-à-vis* the self. The instructions to the subject are often stated in *as if* terms. "You are to take this test as if you were Mr. Jones," or "Pretend you are Jones," etc.

Dymond (1949) published an empathy scale based on this method. Each subject rated (1) self, (2) others, (3) other's self, and (4) other's rating of subject's self, on six traits: self-confidence, superior-inferior, selfish-unselfish, friendly-unfriendly, leader-follower, sense of humor. Deviations between criterion and predictions were computed. The range of scores indicated a dimension of personality. The only test of validity was a comparison of empathy scores with independent judgments of a person's empathic skill on the basis of TAT protocols. The correlations were satisfactory, although as Dymond says, "hardly evidence on which to state that this is a valid test of empathic ability" (Dymond, 1949, p. 131).

Notcutt and Sylva (1951) used a similar procedure in developing a prediction test. Eighteen traits, taken from Murray's system, were scaled and checked by husbands and wives for self and other. The predictions of the other were significantly greater than chance. Successes were greater on traits where subjects made similar self-ratings.

Dymond's and Notcutt and Silva's studies were criticized by Hastorf and Bender (1952) on methodological grounds. They point out that no correction is introduced for "projection." In their own study, using the Allport-Vernon Study of Values, they obtained two scores: an empathy score, which was the difference between the prediction of the other's responses and the criterion (the other's actual test responses); and a projection score, which was the difference between the prediction of the other's responses and the subject's *own* responses. In their study of 50 subjects, 20 had higher empathic scores (predictions were closer to the social objects' actual responses); 28 had higher projection scores (predictions were closer to *own* responses). Two subjects had equal empathy and projection scores. Thus projection was more common than empathy in this group of subjects.

These studies illustrate how empathy — the use of *as if* in social behavior — may be investigated experimentally. In a later section, some additional studies are reviewed which are aimed at the origins of *variability* in taking-the-role-of-the-other. The general thesis is that variability in role perception, role enactment, and taking-the-role-of-the-other is a function of the self. To establish this thesis, we now turn to a psychological analysis of the self.

THE SELF

More than any other single group, the role theorists have developed and used the conception of the self as an intervening variable. The historian of modern social psychology could trace — in an unbroken line — the rise and development of the self from William James, Baldwin, Cooley, and Mead through Faris, Park and Burgess, and Kimball Young to the present. The thesis central to this group of writers was that the self arises out of social interaction. The self was seen as predominantly, if not exclusively, social. The core of Cooley's idea, "Each to each a looking-glass reflects the other that doth pass," is seen in the theory of selfhood advanced by these sociologists.

Among psychologists, Allport, more than any other, has emphasized the centrality of the self in his conception of trait organization. His book (1937) followed by his challenging paper (1943) re-awakened interest in the self among psychologically oriented theorists. This interest was given additional — though not direct — impetus by Lewin (1935) and by such Gestaltists as Koffka (1935). Today, no new treatise or textbook on social psychology is complete without a discussion of the self. (See, for example, Asch, 1952; Hartley and Hartley, 1952; Krech and Crutchfield, 1948.)

Probably as a result of the convergence of normative infant studies and the observations of the psychoanalytic child psychologists, Mead's notion of the self as originating solely in social interaction gave way to the notion that the self has earlier origins. Since the trend in role theory is in the study of the interaction of self and role as coordinates and not as parallels, it behooves us to give a detailed account of the origins, functions, and dimensions of the self.

Borrowing from James, Symonds (1951) regards the self as the object of awareness — the ego as the acting, knowing, perceiving object. Brewster Smith (1950) regards the self as a phenomenal entity, the ego as a construct. The present writer has used the term self to refer to the phenomenal experience of identity, an inference arising from a person's experiences with things, body parts, and other persons. "The referent for the 'I' . . . is the cross section in the present of the total cognitive organization that embraces those more-or-less enduring substructures, the empirical selves" (Sarbin, 1952a, p. 21). The more cumbersome term, self-as-inferred-by-others, has been used to refer to the products of behavior analysis. For convenience, we may substitute the more accessible word *ego* for the latter term. Thus, in the discussion which follows, self refers to a phenomenal experience (not all of which can be communicated via language); ego refers to inferences made by psychologists (or others) about the more central or enduring psychological characteristics of a person.

Neither self nor ego is directly observable, but inferences can be made about both. The person can communicate about the phenomenal experience of selfhood by the use of I-sentences, for example "I am rugged and powerful." The observer can likewise communicate about the ego of a person by the use of third-person sentences in which adjectives, such as dominant, self-confident, original, reverent, etc., are the significant terms. The inferences about ego structure can be derived from phenomenological report or from observations of nonverbal behavior. Whether making inferences about self or ego, the essential terms used stand for *qualities*, rather than actions. These qualities may be conceptualized by observers in several ways. The present author, following Allport (1937), finds the following useful: traits, attitudes (including sentiments), and habits. These are inferences derived from observing actions of persons. The trait is a more general quality than an attitude, which in turn is more general than the habit. Because of space limitations we shall omit a treatment of attitude and habit and, in general, equate quality and trait.

As a cognitive structure, the self develops out of interaction between the organism and stimulus objects and events. In early life, the "motivating conditions" which lead to action arise out of perturbations in the steady state. In the neonate, the steady state is maintained via chemical equilibrium or homeostasis. In the process of maintaining equilibrium, trace residuals are laid down within the nervous system — the earliest cognitions. These trace residuals are integrated within the organism; the "steady state" is a more complex affair than formerly. Homeostasis now serves a more complicated organism, although still in an automatic fashion. The beginning of the perceptual constancies reflect the equilibrating function of the human organism at a still more complex level. Object constancy is an adaptive mechanism — were it not for this phenomenon, a person would continually be questioning the validity of his perceptions as he encountered the same objects under different conditions. In the same progression, images may be subject to the same constancy mechanism as objects. In the face of new information to the contrary, a person's image of an object or person remains relatively constant. Similarly, concepts remain constant under diverse conditions. The constancy of attitudes under planned attempts to change them is a further illustration of constancy of conceptual structures.

The application of the equilibratory principle from chemical homeostasis to conceptual constancy is extended to all cognitive structures. The concept of the self and its substructures are subject to the principle of constancy. This is the theme in Lecky's *Self-Consistency* (1945). At this juncture, suffice it to say that the self as a cognitive structure in general functions according to the constancy principle.

Toward the maintenance of what constant state does this principle operate in early life prior to the apprehension of agents and objects? The answer is in part conjectural, but available evidence suggests that the stable state of early infancy is made up of somesthetic *qualities*. With an immature neuromuscular system, the organism has no way of "knowing" agents or objects. It does perceive — primitively — such qualities as tension and quiescence. Qualities, then, are the first cognitions — or, in the older language, the first contents of consciousness. Before the dawn of gesture or language, residues of qualities are laid down and form the nucleus of the self (Freud, 1935; James, 1890).

The self, then, is seen as an organization of *qualities* at first unverbalized and unverbalizable, later verbalized in part by gestural or linguistic devices such as naming, self-drawings, adjectives, I-sentences, etc. The origins of the self lie in the resultants of the interaction between two series of events. The first is the maturational series; the second, the personal-social series. The *maturational* series of events may be assumed to be fairly constant from culture group to culture group. Within certain limits of variation, the age at which development of the nervous system is complete, the age at which ocular focusing begins, the age at which dentition occurs, etc., is the same for human infants in all societies.

The *personal-social* series of events is the major determinant of psychological variation. Persons enacting roles of mother, nurse, father, sibling, and others interact with the child according to prescriptions and proscriptions characteristic of the culture, subject to the minor variations of individuals. Of particular interest are child-rearing practices and their influence on the development of the self. This has been the focus of study of personality-and-culture in the past decade. Of equal importance in assaying the origins and structure of the self is the *time* at which specific child-rearing practices are introduced or terminated. Here the concepts of maturational series and personal-social series converge. For example, child-rearing practices which lead to prolonged feeding frustration in early infancy will have different effects on the formation of the self than the same kind of frustration introduced when the child is at a more mature physiological level. Oblique evidence for this assertion comes from an attempt at establishing the relationship of *time* of weaning to adult ego structure. Goldman (1950), by means of a 19-trait scale, placed her subjects on an optimism-pessimism continuum. Then she compared the scores with the age at which breast feeding was terminated. This information was obtained from the mothers of the subjects. The results indicated that subjects weaned before five months were rated more pessimistic than subjects weaned later.

Although such studies as Goldman's lend support to the importance of both the maturational and personal-social series in the formation of the self, at the present time most of our information about the effects of timing infant-rearing practices is analogically derived from animal studies (Hunt, 1941; Hunt, *et al.*, 1947).

Werner (1948) has supplied a number of concepts which may be applicable to the study of the self. Among these is the dimension of differentiation as it applies to perception and to motoric behavior. Time binding and tension binding are corollary conceptions with which cognitive phenomena may be described. For example, the newborn infant may be characterized as undifferentiated in action and perception, with no ability to bind time or tension. The mature adult may be characterized as differentiated in actions and perception and, within specified limits, able to tolerate time and tension. These dimensions will be discussed in the description of the empirical selves which go to make up the self structure.

Another conception necessary for an understanding of the self and its place in role theory is that of fixation. The agents of culture, either through intentional training procedures or through unwitting performances, may prolong the organization of the self at a particular level, thus preventing the full development of later structures. Werner (1948) has pointed out how overspecialization at earlier levels inhibits creativity at later levels. Thus if a person were fixated, either through overlearning or trauma, at a developmental level of the primitive-construed self, the development of the social self would be inhibited, and we would expect the person to be incapable of performing certain adult roles which call for time binding, tension binding, and highly differentiated perceptual-motor actions.

The concept "fixation" has been used in many ways. Our usage of the term is intended to convey the meaning of a relatively permanent change in the structure of the organism, which change follows from perturbations introduced under conditions where the organism cannot differentiate cues. The young infant, by virtue of its incomplete maturation, is unable to distinguish the environmental events associated with the experienced quality. The response, in this case a quality (or feeling), is linked to any stimulus event which is in some way equivalent to the initial stimulus event. (Bearing on this notion are the experiments of Stellar and Morgan, 1953, and Fredericson, 1951.)

With these notions: the constancy principle, the maturational and personal-social series of events, the dimensions of cognition, and fixation, the early organization of the self may be described. Elsewhere, the author has presented a more complete analysis of the origins and structure of the self (Sarbin, 1952). A modified presentation is offered here.

As is the case in all organization in nature (Sickles and Hartmann, 1942), residuals resulting from actions designed to maintain constancy are organized around foci. These foci of cognitive organization we shall call empirical selves. The first such focus arises out of the infant's automatic responses to imbalances introduced by the change from the stable intrauterine environment to the unstable extrauterine environment. Unless imbalances in the child's physiology are corrected by the intervention of other persons, the child cannot survive. The efforts of persons engaged in child care are directed almost exclusively toward maintaining the physiological equilibrium by means of temperature control, protection against natural hazards, and maintenance of the food and water supply. We have called this first focus the *somatic self* (S_1), which arises from the infant's experiences during the first few weeks of life when the primary concern of the mother or nurse is the maintenance of the infant's *soma*. This inchoate cognitive structure lays the groundwork for differentiation of the self from nonself.

At the beginning of the development of this first structure, the infant does not distinguish between external objects and the various parts of its own body. Differentiating its own body from the external world — establishing the boundaries of the somatic self — is one of the chief activities of early infancy. The infant comes to recognize that different sensations result when it strikes objects in the external world or when it pinches itself. Single stimulation occurs in the former instance, double stimulation in the latter. The *Anlage* of the postural model of the body is laid down as the result of differentiating the effects of this differential stimulation. This is the cognitive structure which later develops into the body image (Schilder, 1935). Overdevelopment of this structure through trauma or through prolonged attention to purely somatic disequilibrium may lead to an overspecialization or differentiation of this structure. If this occurs, it is likely that pathological behavior will result — behavior in which the body image is central. Severe hypochondrial behavior, narcissism, and somatic delusions are examples of adult conduct where the body image or somatic self is overdeveloped.

Not all imbalances are immediately corrected by the child-caring persons. Uncorrected homeostatic imbalances lead to tensions in the organism. These tensions, such as follow from hunger, thirst, accumulation of waste products, etc., become part of the stimulus field. Receptor and effector processes lead to tension reduction. The process appears to be similar to "drive discrimination." Of course, this means that certain maturational events have occurred so that the receptor or effector structures are functional. By age three or four months the infant can respond to external stimuli, although in an undifferentiated way. When a child is hungry and is given food, the giver can be perceived only as the agent of tension reduction, not as an object or person.

With the maturation of the organism and with its increased motility, other kinds of stimuli are mediated by means of the receptor organs. At this level of self-organization certain events, either external or internal, become associated with tension reduction.

The main characteristic of the organism at this time is the differentiation of qualities within the organism. These qualities are probably bipolar: a quality which we designate *tension* and a quality which we designate as *quiescence*. The operation of the homeostatic mechanism is aided by the actions of the child-rearing agents, the result of these actions being tension reduction. We can infer only that the child perceives the tension and quiescence as diffuse generalized qualities, probably the equivalent of "feelings." Chronologically, the infant is under five or six months of age when this process begins. There is some casual evidence that prolonged deprivation leads to fixation of qualities, which in turn direct the human organism's activities towards tension-reduction processes. DuBois' analysis of the Alorese (1941) lends credence to this notion. The quality of tension associated with prolonged hunger during this phase of development "fixates" the self-organization around the receptor-effector focus. The specialized attitudes toward food and hunger may be considered the residue of this overspecialization.

Since the receptor and effector processes are the focus of the child's rather undifferentiated self-organization, we have labeled this second focus the *receptor-effector* self (abbreviated S_2). Around this matrix the infant at the primitive deductive level (conditioned response) interacts only with that nonself which is instrumental in reducing or eliminating tensions. No time or tension binding is apparent. Perception is syncretic, actions are mainly of an undifferentiated character.

The third level of development we have called the *primitive construed self* (S_3). The beginning of this focus of cognitive development is about six months of age. Neuromuscular maturation has progressed to the point where the infant has begun to acquire some control over his movements and to express his needs through motor activity. He is no longer an entirely passive recipient of the tension-reducing activities of nonself agents. His increased motoric activities bring him into contact with a larger number of stimulus objects. Perceptual development has also become more differentiated, and the infant can perceive differences between *agents* instrumental in tension reduction and *objects* instrumental in tension reduction. It is doubtful whether the child can perceive differences between objects and persons *qua* persons. Bayley (1933) has shown that between the ages of four and six months differential vocalization to nonself objects occurs. After six months of age differential vocalization to nonself *persons* begins (Shirley, 1933).

Motoric development is such that the child can express pleasure, eagerness, and displeasure. To a limited extent he can discriminate tone and expression. The criterion by which a child distinguishes objects from persons is the ever-changing nature of the stimulus properties of things. The focus of cognitive development here is the laying down of traces of perceptions of persons who are selective, dynamic, and somewhat unpredictable. Baldwin described the process as follows: ". . . this is the child's very first step toward a sense of the qualities which distinguish persons. The sense of uncertainty grows stronger and stronger in its dealings with persons. A person stands for a group of experiences quite unstable in its prophetic as it is in its historical meaning. This we may, . . . in assuming it to be the first in order of development, call the 'projective stage' in the growth of the child's consciousness" (Baldwin, 1902, p. 14). The child's cognitive organization must now be able to select the more enduring stimulus qualities from the more evanescent.

Tension binding has not yet begun. Available evidence indicates that perception of stimulus qualities and motoric discharge are continuous processes. No vicarious trial and error is observed. Because of the immature neuromuscular development, the infant can no more than cry or show diffuse restlessness if tensions are not quickly reduced by other persons. Now that the infant has some control over its muscular activities, he can *act out* his responses immediately. Whether the stimulus conditions are external, as in the case of a brightly colored object, or internal, as in the case of hunger, the infant responds *immediately* by reaching, grasping, shouting, crawling toward, etc. The young infant is still incapable of binding time or binding tension. The ability to wait, to tolerate delay between stimulus and response, does not develop until the beginnings of the next level of selfhood. Extrapolating from available data on delayed responses in children, it is probable that tolerance of delay at this stage is not more than a few seconds. In the child of ten or eleven months, delay, followed by responses indicating memory, is not more than one minute. At this level the infant's behavior is less syncretic than before; he can make discrete responses to specific stimuli selected from the undifferentiated mass of stimulation received through the several sense modalities.

At this point we infer that the child has three foci of cognitive organization: the somatic self, the receptor-effector self, and the primitive-construed self. These are in the nature of reference schemata. They serve as background against which incoming stimuli are perceived and responded to. They are trace residuals, and are derived from interaction with agents, objects, and persons, the interactions originating in the preservation of constancy.

In the fourth quarter of the first year the beginnings of a new cognitive structure may be inferred from the child's behavior. The focus is the gesture (Mead, 1934). With increasing experiences and further maturation the child can associate gestures with things and with persons. Now he can discriminate clearly between persons and objects. The substructures so far described permit the development of a new emerging self which we have called the *introjecting-extrojecting self* (S_4). The development of motor skills, increased motility, postural changes, prehension, etc., increase the number of nonself objects and persons that can be responded to and organized into the currently forming empirical self. Words are differentiated and imitated, and the child can actively participate in play.

Many observers of young children have described the development of linguistic behavior in relation to the self. Space limitations prevent discussion of the contributions of Piaget, Cooley, Lewis, and others. With language symbols the acts of others can be grouped into a concept "similar to my own" (extrojection or projection). Similarly, one's own acts can be grouped into a concept "similar to others" (introjection or identification). This reminds us of Baldwin's "dialectic of personal growth" (1902). He points out how the reactions of an infant to the shifting stimulus properties of others become assimilated or "subjective" if congruent with the mental structures so far developed. The corollary process, "ejective consciousness," occurs at about the same time. This is the same as extrojection and appears to result from a primitive analogical process which depends upon the use of language.

The child's cognitive development is such that now he can differentiate not only agents, objects, and persons, but he can differentiate the *acts* of persons. Functionally adaptive responses to the differential acts of others become the order of the day. At first, the responses of others are perceived as approval-disapproval. Here we have the prototype of learning by reinforcement. We also observe the beginnings of tension binding and time binding. The child operating under the aegis of approval-disapproval learns to withhold or inhibit certain motoric responses upon the perception of internal or external stimuli. The discrimination between approved acts and disapproved acts is the foundation of the sense of shame. His ability to form concepts is limited so that the acts of others cannot be organized into roles and, of course, he cannot yet take the role of the other.

The fact that the child can cognize acts of others and of self into groupings reflects the beginning of the *socius*. Fixation at this level leads to behavior which is primarily "other" oriented. That is to say, actions must first be considered in terms of possible approval or disapproval, reward or punishment. Extensive development of this aspect of behavior, of this empirical self, may inhibit the later development of the role-taking ability. Excessive time binding and the perception of *discrete acts* rather than *organized wholes* prevent the more adaptive, relatively rapid taking-the-role-of-the-other.

With the increase in language development, particularly the use of the pronoun "I," the child attains a more refined concept than before, at the same time acquiring a conventional symbol for self-reference. Before this, in order to communicate about himself, the child was restricted to motoric signs or names which were used as others use them. With the development of the "I" concept, which includes my, mine, me, and myself, the child can quickly and easily differentiate self from others. His language development also allows for the grouping of the actions of other persons into social roles. Certain actions are associated with "mother"; other actions are organized into the concept "brother." The importance of linguistic phenomena in concept formation needs no further emphasis here save to say that, cognitively speaking, roles are concepts. Here the *social self* (S_5) emerges, and its differentiations are in terms of roles rather than discrete acts as in the case of the just preceding empirical self-structure.

Taking-the-role-of-the-other now becomes a part of the perceptual and motor behavior of the child. The incorporation of "other" roles can readily be seen in his play activities. He can participate in playing "let's pretend," the first effective test of the *as if* skill.

Linguistic devices are the tools which aid in the further development of the social self. The learning of language symbols, particularly those that have social referents, aids in the perceiving, identifying, and conceptualizing of roles and role expectations.

The child can now assign to the self (the referent for the "I") the same terms used by others. The assignment of qualitative terms, such as good, bad, dirty, unkind, stupid, etc., are the "reflected appraisals" of others. A theoretical account of this development is given by Kimball Young (1945, 1952), based on Mead's writings. [The work of Piaget (1926) should also be mentioned for its important observational descriptions of the attainment of concepts.] In summarizing Mead's viewpoint, Cottrell and Gallagher (1941) state:

> If we hold to Mead's view that meaning inheres in the ability to call out implicitly in one's self the role which one is calling out explicitly in the response of another, we can shift our attention from the philosophical aspects of the process upon which he concentrated, and focus instead upon the experiences that limit and frustrate this role-taking ability. (p. 27)

It is to these "experiences that limit and frustrate this role-taking ability" that research should be directed. Below are two studies which illustrate the effects of different kinds of experience on the social self — one conducted in another context, the other within the context of role theory.

It is well established that the child with adequate symbolic development is capable of forming concepts. If the concept of self is the reflection of the evaluative actions of others, then we would expect to find differences in self-concepts among persons who experience different kinds of socialization behaviors. Sears (1950), for example, has shown the influence of different kinds of maternal child-rearing actions on the ego structures of first and second children. He reports the results of several investigations which converge to show a positive correlation between degree of frustration in nursing and weaning and dependency behavior during the preschool period. Dependency behavior as studied and defined is a quality of the more central aspects of the person inferred by an observer, hence it may be labeled an aspect of the ego structure.

Returning to the constancy principle, we have already mentioned how a concept, whether of a physical object, a social object, or of the self, tends to remain constant. The structure of the self develops out of experiences with events, objects, and persons. If the further development of the social self is a result of the "looking-glass" principle, then we would expect different kinds of self-concepts in persons whose early socializing interactions were limited to one or two parent figures from those whose early socializing interactions were more extensive. The more parent figures in the socializing process, the more likely that the common qualitative aspects of the culture will be learned.

This leads to the testable hypothesis that the person whose early socialization is mediated by reciprocal role taking with *many* significant persons will have a different self-concept from the person whose early socialization was the responsibility of one or two parent figures. Further, those with multiple parent arrangements will be more socially adjusted than the others and will share many common traits. As a class exercise, 90 upper division students (30 males, 60 females) were asked to describe the self on a 200-word checklist. They were also required to furnish certain personal data, among these the number of persons responsible for their care and rearing during the first five years. It is important to note that the criterion for the number of parent figures was left to the discretion of the subjects and hence was as phenomenological as the criterion for determining which of the 200 adjectives were self-descriptive.

The distribution of subjects in regard to the number of parent figures was as follows: one or two parent figures, 54 (12 males, 42 females); 3 or more parent figures, 36 (18 males, 18 females). For each of the four groups thus derived, the number of checks for each adjective was counted. This was converted into a proportion. Where the difference in proportion between the two groups of the same sex was at least .20, the adjective was retained as characteristic of the self-conceptions of members of that group.

Although the groups with three or more parent figures *did not check* a significantly larger number of adjectives, they did check more of the *same* adjectives, indicating that those who have more significant "others" in their childhood tend to have more similar self-concepts.

When we examine the quality of the adjectives characteristic of the groups, we find a further demonstration of the influence of the number of parent figures on the self-concept. Both men and women of the one- or two-parent figure classification describe themselves as restricted and nonsocial. The members of the other classification describe the self in a number of adjective clusters: sociable, logical, imaginative, affectionate, and argumentative. These tentative conclusions are warranted: persons reared by multiple-parent figures have self-conceptions that are: (a) more "social" and extended, and (b) more homogeneous than persons reared by one- or two-parent figures (Sarbin and Van Spanckeren, 1953).

Space considerations limit exposition of the further development of the social self. The reader should not interpret the choice of illustrative material to mean that postinfantile events have no influence on the development of the self. However, the self, as organized in the earlier years of life along the dimensions described before, serves a selective and directive function, thereby placing limits on the number and quality of role perceptions and role enactments.

EMPIRICAL METHODOLOGY FOR ASSESSING THE SELF

Although the foregoing is by no means the final word on the theory of self-development, it reflects the tenor of the times. Within the past few years such writers as Rogers (1947), Hilgard (1949), and Newcomb (1950a, b) have helped re-establish the concept of the self. With this re-establishment has appeared the necessity for discovering methods for assessing the self.

Recapitulating, the self is what the person "is," the role is what the person "does." When interested in the self, we regard the person as an organization of *qualities*. When we study roles, we regard the person as an organization of *acts*. Parenthetically, direct observation reveals only action systems resulting from the interplay of self and role.

The quality is not something that is directly observed, as in the case of the act. The quality must be inferred from acts. In English syntax, the usual form for communicating qualities is the adjective. Thus the adjective is a symbol which can be used in describing a quality of the self. At first glance, this seems too simple to describe the complexities of the self-structure. Oversimplified as everyday language appears to be for describing a phenomenal datum, it is, as Allport says (apropos of trait names):

> . . . several grades more adequate than the mathematical symbols and neologisms that psychologists sometimes employ. The nature of our problem forces us to seek out, to identify dynamic mental structures and substructures, and *to name them*. And this is necessary even though the lexicon of any language is far from offering a perfect catalogue of the elements of mental life . . . (Allport, 1937, p. 310)

If we are interested in phenomenological reports, of course, we can record a person's autobiographical statements about the self. The uneconomical nature of this method for purposes other than therapy requires no comment here. This method may be abbreviated by the use of I-statements, such as in personality inventories or in attitude questionnaires. Often, however, the units of such statements confuse *acts* with *qualities*. A direct and unconfusing approach to qualitative description of the self is the use of an adjective checklist (Gough, 1950).

To illustrate the method and its advantages, Sarbin and Rosenberg (1952) performed three experiments. In the first, the self-conceptions of diagnosed neurotics and normal controls were compared. In the second, the self-conceptions of men and women students were compared. In the third, the self-conceptions of three religious groups were compared. Only the findings of the first experiment will be reviewed here. At least 70 percent of the neurotic group checked the following qualities: anxious, changeable, confused, cooperative, dissatisfied, emotional, gentle, imaginative, immature, nervous, self-centered, worrying. At least 70 percent of the controls checked these qualities: active, adaptable, affectionate, civilized, conscientious, considerate, curious, dependable, fair-minded, friendly, intelligent, sincere, sociable, tolerant.

If the postulate is accepted that the self is cognized as an organization of qualities and adjectives are linguistic devices that enable persons to communicate about qualities, then the adjective-checking method as illustrated here has an important place in social psychological research.

Another promising method for getting at the phenomenal experience of self is the W-A-Y technique reported by Bugental and Zelen (1950). The respondents are simply asked "Who are you?" three times in succession. The three responses are then classified according to 17 categories such as age, sex, social status, name, occupation, etc. Differences in self-conception are noted for objectively different groups; e.g., older subjects conceptualize the self differently from younger subjects in the use of age and social status categories.

The authors advance some tentative conclusions as a result of their preliminary investigations. Among these conclusions are: the frequent use of the "name" category suggests the name as a central feature of the self; the personal pronoun category was used more frequently by women and by older persons; the sex category was more frequently used by younger than by older subjects — the authors suggest the possible relationship between sex identification and psychosexual development; the age category differentiated among the groups — it was used most frequently by persons who were "at the apparent peak of physical and social abilities" and less frequently by younger and older persons.

In the review of experiments emerging from the theory, still other methods will be described. We should make clear that this is a representative but by no means complete catalog of empirical methods for assessing the self.

SOME EXPERIMENTS EMERGING FROM ROLE THEORY

The worth of any theory lies in its potential for suggesting hypotheses. The pragmatic value of role theory is indicated by the kind and number of hypotheses generated by it which are subject to empirical investigation. Role theory, with its emphasis on the self, has stimulated a number of researches, some of which are reviewed below.

TENSION BINDING

One example of a study stimulated by this theory of the self is concerned with the dimension of tension binding. Simply stated, tension binding is the ability to inhibit a motoric act once the stimulus to the act has been perceived. The dependence of tension binding and time binding on the development of *as if* behavior has already been noted. The ability to bind tension makes it possible for a person to inhibit inappropriate role behaviors. The dimension may be thought of in quantitative terms — certain people can inhibit motoric acts for long periods of time, others can tolerate the tension only for brief periods. It has become customary to refer to those who are deficient in this ability to bind tension as undercontrollers, that is, as lacking in self-control; and to those who bind tension more than objective conditions warrant as overcontrollers. Clinically, persons diagnosed as psychopaths exhibit the inability to bind tension; they act on impulse; they are unconcerned about the approval or disapproval of others. Persons diagnosed as neurotic characteristically overbind tension; they act compulsively rather than impulsively; they are overconcerned about the approval or disapproval of others. The overt conduct of these classes of persons indicates some directing or selective organization at work.

We consider this organization to be the self. The undercontrollers, those who cannot bind tension, are presumably characterized by a self-structure where the primitive construed self (S_3) is the dominant substructure. The overcontrollers, those who overbind tension, are presumably characterized by a self-structure where the introjecting-extrojecting self (S_4) is the dominant structure. The characteristics of these self-structures as defined allow for the prediction of role behavior under various conditions.

The tendency to act immediately upon the presentation of an annoying stimulus may be fulfilled or inhibited. In organized social living those tendencies to act are constantly being brought to the surface by multifarious annoying stimuli. Persons with a low threshold for binding tension would respond more frequently or more intensively to annoying stimuli, while persons who have a high threshold for tension binding would respond less frequently or less intensively to the same stimuli. Furthermore, the state of being annoyed is more easily attained by persons who cannot delay their responses, during which delay vicarious trial-and-error (VTE) (Tolman, 1939) behavior is possible, the content of the VTE's being questions of approval or disapproval by others. Conversely, the state of being annoyed is less readily attained by persons who can delay responses, during which delay they can vicariously assess whether the acts in question will gain the approval or disapproval of others. Such delay is frequently nonadaptive, in that the moment for adaptive action may have passed while the extensive VTE behavior is occurring. (This is the essence of clinically defined neurosis, particularly the compulsive type.) Between these two classes of persons are those who may be characterized as "socially adequate tension binders" (S_5). These are persons who can bind tension according to objective demands, who require less time for VTE behavior because of their use of an organized role-taking approach. That is to say, the quick and efficient grouping of stimuli can be performed with the aid of a concept — in this case a role — rather than the more time-consuming approach of applying an approval-disapproval criterion to every incipient response.

Gough's role-taking theory of psychopathy (1948) (after Mead, 1934) embraces some of the same ideas. He asserts, after a review of the literature, that the psychopath is fixated in such a manner that he cannot perceive the self as an object; he cannot take another person's perspective. This aspect of role taking, emphasized by Gough, is substantially the same as our tension-binding dimension. When frustrated, the psychopath cannot take the time to put himself in the role of the other. He cannot maintain the *as if* set. Once an act is begun, it is carried through to completion.

In order to test this hypothesis that psychopaths, or poor tension binders (S_3), would be the most responsive to annoying stimuli, we administered a 172-item annoyance questionnaire, developed by Harsh (1936), to 60 male subjects: 20 diagnosed as psychopaths (S_3), 27 diagnosed as neurotic and representing S_4, and 13 without psychiatric diagnosis (normals) representing S_5. The annoyance inventory was divided into thirteen categories: e.g., sensory annoyances, inconsiderateness of others, matters of dress, etc. The means of the three groups were compared on each of the category scores and on total scores. On total mean scores, differences between means of psychopaths and normals were significant ($P = .01$); between normals and neurotics $P = .001$. In twelve of the thirteen categories, the magnitude of the scores was in the predicted direction, the psychopaths had the highest mean annoyance score, the neurotics the lowest, and the normals intermediate (Albrecht and Sarbin, 1952).

SKILL IN TAKING-THE-ROLE-OF-THE-OTHER

Taking-the-role-of-the-other is one of the qualities absent in persons fixated at the level of the primitive-construed self. If the dominant cognitive structure in the psychopath is this empirical self (S_3), then we would predict poor scores on tasks involving the participation of the social self. The most obvious kind of performance to test this hypothesis is an empathy test such as described earlier in this chapter.

Baker (1954) performed an experiment on a group of male prisoners in an effort to test the hypothesis that psychopaths would be less able to empathize with their cellmates than nonpsychopaths. All the subjects had shared cells with their cellmates at least four weeks. Assignment to the category of psychopath was affected by the use of three MMPI criteria. The procedures were carried out in one sitting. Each subject and his cellmate filled out four adjective checklists: (A) for self, (B) for cellmate, (C) prediction of self-checks of cellmate, and (D) prediction of (B), subject's guesses of how he appears to his cellmate. The predictions made by each subject on checklists (C) and (D) were compared with the cellmate's actual choices on checklists (A) and (B). These comparisons provide two objective measures of empathy: the first, percent correct predictions on list (C), the second, percent correct predictions on list (D). (C)-percent in part gets at the ability to perceive traits or qualities that a person assigns to the self; (D)-percent in part gets at the ability to perceive traits or qualities others use in assessing one's self.

Although the samples are small, 21 psychopaths and 13 nonpsychopaths, the differences on (C)-percent and on (D)-percent are significant. The reduction in statistical reliability introduced by a correction for subject-object similarity did not appreciably alter the findings. Small positive correlations between empathy tests and age and education were also found. As a whole, the experiment devised as a result of hypotheses arising from role-taking theory verified the predictions.

Cline (1953) has reported some findings which tend to confirm hypotheses arising from this theory. By means of a hidden camera and microphone, he filmed a number of "job applicants" in a realistic life situation. These "job applicants" will be referred to as *social objects* in the following discussion. Seven of the films (11 minutes each) were used in his study of the assessment of judges of personality. These films were shown to the subjects of the study, who were required to check an adjective checklist as they thought the social objects had checked it. This is one test of taking-the-role-of-the-other. A second test involved making postdictions about the social objects. For this test, Cline made up a series of multiple-choice items on the basis of extensive interviews with parents, friends, and associates of the social objects. Only items were used upon which there was consensual agreement. Here is an example: When the subject (seen in the film) is in a violent argument he usually (a) becomes very sarcastic, (b) uses profanity and obscene words, (c) leaves the room (or area), (d) strikes his opponent with his fist.

Using a judging index made up of these two tests, which appear to test the ability to take-the-role-of-the-other, Cline compared the subjects in the top and bottom quartiles on a number of variables. Among these were scales of the MMPI, the California Public Opinion Survey, and the Henmon-Nelson intelligence scale.

From our theoretical emphasis on the importance of early fixation, the prediction would follow that a person in whose cognitive organization the somatic self (S_1) is a dominant structure would be less skillful in role taking

(S_5). Hypochondriasis is a syndrome in which the somatic self dominates all the other selves. In Cline's analysis, the hypochondriasis scale of the MMPI, at a level of confidence better than .01, differentiates those scoring high on the role-taking tests from those scoring low. The hypochondriacs get lower scores on the role-taking tests.

If the assumption that certain kinds of psychopathology follow from fixation of the self at different maturational levels (Sarbin, 1952) is granted, then adult schizophrenia is seen to be related to the dominance of the receptor-effector self (S_2) and psychopathy to the primitive-construed self (S_3). Persons scoring high on scales which tap schizophrenic or psychopathic behavior, then, should do poorly on role-taking tests which sample behaviors mediated by the social self (S_5). Cline's results confirm the theoretical predictions; the schizophrenic scale and the psychopathic-deviate scale of the MMPI, at levels of confidence between .01 and .05, differentiate the top and bottom quartiles on the role-taking tests in the predicted direction.

Some other variables which differentiate the top and bottom quartiles consistently with the theory are: intellectual efficiency (MMPI), social status (MMPI), prejudice (MMPI), and intelligence. Ethnocentrism and Authoritarianism (E and F scales of the California Public Opinion Study) likewise show differences in the predicted direction.

An experiment by Sarbin and Stephenson (1952) was also aimed at identifying the characteristics of good role takers. The hypothesis stated that persons characterized as rigid would be less able to adopt a role which was incongruent with the ego structure. Persons who were classified as flexible *would* be able to adopt a role not congruent with the ego structure.

On the basis of Rokeach's experiment (1948) rigidity and flexibility were defined as high and low scores on the ethnocentrism scale. The procedure was as follows: Two undergraduate classes in psychology provided the subjects (24 men, 28 women). The University of California Public Opinion Survey was administered (this includes a political-economic conservatism scale, an ethnocentrism scale, and an authoritarianism scale). One week later each member of the class filled out the authoritarianism scale "as you think a person named Mr. (Miss) Smith would take it." Smith is described as an intensely prejudiced person. Two weeks later the same procedure was followed except that the role instructions were now of Jones, an equalitarian person. Our predictions were as follows: persons scoring high on ethnocentrism would more readily enact the congruent authoritarian role than the incongruent equalitarian role; persons scoring low on ethnocentrism would be able to enact both the incongruent authoritarian and the congruent equalitarian roles. The analysis confirmed our predictions. Correlations between scores on ethnocentrism (without role instructions) and authoritarianism *with* the role instruction to play the part of Smith, the authoritarian person, were —.29 for males and —.41 for females. The flexible subjects (low ethnocentrism scorers) were better able to change their responses. The correlations between the same two scales when the role instruction was to play the part of Jones (the equalitarian person) were .65 for males and .78 for females. The rigid persons (high scorers on ethnocentrism) did not change their responses under the role instructions in the direction of equalitarianism. Taking-the-role-of-the-other, an aptitude or skill, and equalitarianism, a cluster of attitudes, appear to be related.

This experiment aimed at ego-structure variables yielded as a by-product a procedure for getting at the phenomenal self. *All* the subjects regarded the "equalitarian attitude" as a quality of the self. That is to say, despite their checking items characteristic of a fairly wide spread on the authoritarianism scale, few subjects checked different items when asked to take the role of the nonethnocentric person. For example, if a subject's score on the first test was high (ethnocentric) when taking the equalitarian role, the score remained high. The inference from this incidental finding is that a discrepancy exists between conception of self as a phenomenal object and the ego (inferences made about the self by others). From this finding and from the usual definition of insight as consistency of self-evaluation with an outside criterion (usually the pooled judgments of others), the tentative statement follows that ethnocentric persons are deficient in insight (cf. Adorno, *et al.*, 1950). If we had used a suggestion of Hastorf and Bender (1952) we would have found that the ethnocentrics were "projectors," and the equalitarians were "empathizers."

The general picture of the empathic person and the nonempathic person developed from these studies is in line with the tentative descriptions reported by Cottrell and Dymond (1949). TAT protocols were analyzed for high or low scorers on the empathy test. Those high on empathy were judged as emotionally expressive, insightful, outgoing, optimistic, warm, interested in others, flexible, with satisfactory family relations in childhood. Those who scored low on the empathy test were judged as rigid, introverted, noninsightful, inhibited emotionally, subject to emotional outbursts; these subjects experienced difficulty in interpersonal relations, mistrusted others, were less well integrated and had had unsatisfactory family relations in childhood.

Another experiment which points to the pragmatic value of role-taking theory was conducted by Gough and Peterson (1952). The objective of their study was to develop a personality test which would yield predictions of antisocial behavior. This objective was a significant one in view of the generally negative results of previous attempts to differentiate criminals from noncriminals by psychological test procedures (Schuessler and Cressey, 1950). Among other things, the role-taking theory of psychopathy (Gough, 1948) suggested items whose content had no obvious relationship to criminal acts but which was associated with a self-organization characterized by insensitivity to the behavior of others, lack of concern over interpersonal relationships, shallow affect, and inability to take-the-role-of-the-other (syndrome of psychopathy). Like the MMPI, Gough's scale is made up predominantly of I-statements which enable the investigator to make inferences about the ego structure.

The scale (64 items) has been administered to reformatory inmates, Army stockade prisoners, behavior-problem adolescents and appropriate controls. Cross-validational procedures have demonstrated the effectiveness of the scale in separating known delinquents and criminals from normals.

THE SELF AND THE PERFORMANCE OF COGNITIVE TASKS

Using the same typology as in the study of annoyances cited before, Mathews, Hardyck, and Sarbin (1953) investigated the influence of self-organization in the performance of selected cognitive tasks. Predictions were made as follows: Persons about whom inferences could be made that the primitive-construed self was relatively dominant would take a set to "act quickly" when presented with tasks which have both speed and accuracy instructions (S_3). Persons whose self-structures were classified as predominantly introjecting-extrojecting would take the set "act cautiously" (S_4). Persons whose self-structures were classified as social would be more flexible and would approach each cognitive problem with a set pertinent to the task (S_5).

The subjects were 31 volunteer students classified on the basis of MMPI profiles. If the psychopathic deviate and hypomanic scales were elevated more than one sigma from the mean, the subject was placed in the S_3 category; if the depression and psychasthenia scales were elevated more than one sigma from the mean, he was placed in the S_4 category; if all scores save one were within a sigma of the mean, the subject was placed in the S_5 (normal) category.

Among the tests used was the Sears-Hovland motor-conflict apparatus. The reactions of the subjects on the conflict trials were recorded as (1) responding to both stimuli, a "double" response, (2) responding to one of the two conflicting stimuli, "single response," (3) "compromise" between the two stimuli, and (4) "blocking" — no action. We predicted that the S_3 group would be dominated by the action set "*do* something quickly"; the members of the S_4 group "wait, think it over, be cautious" and the members of the S_5 group, variable. A chi-square test significant between .02 and .01 indicated that our groups behaved differently on this test. The S_3's gave predominantly "double" responses (excessively motoric); the S_4's gave "block" responses (VTE's) most frequently, but they also gave a large number of "single" responses. The S_5's gave an approximately equal number of "single," "double," and "block" responses. In general, the predictions were verified: S_3's were excessively motoric, S_4's excessively "bound," and S_5's variable.

Thus, predictions of certain cognitive performances were made from the theory and verified by experiment. The implication of these findings for role theory is that the quality of role enactment, even in the role of a laboratory subject, is a function of the self.

INTERACTION OF SELF AND ROLE

In the first part of this chapter we dealt with concepts in which role was central. Insufficient to account for complex social psychological behavior by itself, the conception of role was supplemented with the conception of self. These are coordinates. Implied throughout the description of each of these ideas is the *interaction* of self and role. Despite the tyranny of literary exposition which forced us to consider but one notion at a time, we have incidentally demonstrated certain of the interrelations of self and role. In this section we deal specifically with some of these interactions and their effects. Of the many questions which could be asked, only the following will be treated.

1. Since the self is a product of experience, what changes in the self follow from the prolonged enactment of a specific role or role cluster?

2. What variations in role enactment are attributable to variations in the self?

3. What happens to the self and/or the role when the person is placed in a position in which the role expectations of others call for performances which are incongruent with the self (self-role conflicts)? A related question is: What happens when the person occupies two positions concurrently and the role expectations of each are incongruent one with the other (role-role conflicts)?

4. Can the independent effects of the self-and-role interaction be isolated? Or, stated in another way: Can self and role as independent concepts be demonstrated empirically?

INFLUENCE OF PROLONGED ROLE ENACTMENT ON THE SELF.

The emphasis in the preceding sections has been, perhaps unwittingly, on early development of self and role. Although the foundation for the self is laid down early in life, this does not mean that residues of experience resulting from adult role enactment have no effect on the self-organization. Notwithstanding, we maintain our historical emphasis: the early form of the structure has selective and directive properties. The effect of prolonged role enactment on the self would be different for, say, the person fixated at the level of the primitive construed self (S_3), and the person whose central empirical self-structure was the social self (S_5). If we take the psychopath as representative of the S_3-fixated person, we would predict little effect on the self-structure of his adult role enactments because of limited *as if* skills.

Merton (1940) has written a logical analysis of a problem which bears directly on the subject under discussion. In analyzing the social structure of bureaucracy, he points up the apparent depersonalization which occurs as a result of prolonged occupancy of a bureaucratic position. This depersonalization probably results from the necessity of regarding clients and personnel as objects. All must be treated according to their classification in the codified rules and regulations; there is no flexibility allowed for personal sentiments or extenuating circumstances. A quality of the self, rigidity, appears to develop out of the daily and exclusive practice in enacting roles which are clearly defined, regularized, and conforming.

One of the hypotheses suggested by this discussion may be stated for later testing: The effects on the self of prolonged occupation of a bureaucratic position are such that rigidity and depersonalization become qualities that can be inferred from behavior in other (non-occupational) roles. The same hypothesis, incidentally, may be stated about any other position in the social structure. Waller (1932) has advanced some hypotheses arising from his analysis of what "teaching does to the teacher." Such hypotheses can best be conceptualized as effects of role enactment on the self.

VARIATIONS IN ROLE ENACTMENT AND THE SELF.

An experiment on self-and-role interaction. The conduct of subjects in an age-regression experiment in hypnosis is plausibly explained by the use of this self-and-role formulation. In the usual experiment the hypnotized adult subject is instructed to enact the role of a child. The performances of many such subjects are indeed convincing. The theoretical question becomes: Is hypnotic age regression an example of play acting (little organismic involvement) or genuine *as if* behavior? The answer, of course, depends on the kinds of tests used to determine the degree of organismic involvement in the *as if* behavior. A crucial test would be one in which the subject performed role behaviors about which he could have no prior knowledge. The Rorschach Inkblots fulfilled this requirement.

The subject could not know from prior experience the kind of responses characteristic of a three- or six-year-old child. Therefore, if the subject, when enacting the role of a three-year-old, reports perceptions which in fact are characteristic of three-year-olds, then we can say that the *as if* behavior is not of the play-acting variety. Sarbin and Farberow (1952) report an experiment in which adult subjects, able to engage in the *as if* behavior of hypnosis, were "regressed" to various childhood levels. The results of the experiment proper were equivocal; some of the subjects seemed to enact the role of the child with remarkable fidelity, as judged from test performance on the Draw-A-Person and from handwriting, speech, posture, and other samples of expressive behavior. At the same time, the Rorschach performances were incompatible with the assigned role. Other subjects did not exhibit the same proficiency in overt role enactment, yet their Rorschach performances were in the range of the assigned age role. The apparent lack of consistency could be explained most readily with the aid of the concepts of self-perception and role perception as inferred from role enactment.

The Rorschach test lends itself to the task of assessing self- and role perceptions. The more enduring ego (matrix of self-perceptions) is *inferred* by the experimenters from the structural features of the protocols, such as kind and number of movement, color, shading, whole, detail, etc., responses. The variation in quantity and time of appearance of these components reflects the level of selfhood. These components are called "validating" or ego factors. They are inferred by the experimenter from the reported object perceptions.

Role perceptions, on the other hand, may be inferred from the role behaviors as seen in the *content* of the Rorschach, (e.g., "It looks like a bunny-rabbit"), from verbal and vocal characteristics, and from mannerisms and gestural responses. The Draw-A-Person test also serves as a test of role enactment. "These expressive features of behavior are the vehicles of communication; it is by the use of such expressive behaviors that the person makes known to others his current role. We shall call these aspects of [role] behavior expressive or peripheral factors" (pp. 119-120).

The performances of two subjects were used as illustrations of the thesis of the interaction of self and role. Thus, one subject's ego is relatively labile — the validating features of her Rorschach performances were actually like those of average three- and six-year-old children. We would infer, then, that her self-perceptions *shifted* with the requirements of each role. Her role enactments, however, were deficient as evidenced by the various expressive measures. Clinical history revealed a relatively immature person who had not yet solved the adolescent problem of independence from mother. "Her self-perceptions were those appropriate to a person occupying the position of a young child in our society. . . . She was deficient in enacting the role, . . . probably as a result of the fixation of the self at an immature level, thus allowing for fewer opportunities for moving through the usual sequence of social positions with the accompanying role behavior. We can infer, then, that this subject's ego is relatively unstable (or labile) and her role perceptions are relatively invariant" (Sarbin and Farberow, pp. 123-124).

The second illustrative subject showed a different pattern of self- and role perceptions. The validating feature of the Rorschach shows a relatively *stable* self-concept, in spite of the different role behaviors performed. Unlike the other subject, she had solved the problem of autonomy. "She has been married and leads an active social life. Under the various role assignments, the validating features changed but little. The expressive features, however, showed gross changes as she moved from one age role to another. The drawing test results, the content analysis, the childlike postures, etc., revealed flexibility in role enactment" (p. 124).

The data seem to make most sense when placed in a system of coordinates (Fig. 2). Role enactment is one axis, the termini of which are flexibility and rigidity. The self is the other axis, the termini of which are stability and lability.

In the two illustrative cases, the first subject's conduct is seen as the result of the interaction of relatively rigid role enactments with a labile self-concept; the second subject's responses are the product of the interaction of flexible role enactments with a stable self-concept. The inference does not follow that these continua are reciprocal. Another subject (x) showed a labile self-concept and a marked flexibility of role enactment. (This subject was later diagnosed as a psychopathic delinquent.) The fourth quadrant (stable self, rigid role enactment) would contain such persons as compulsive neurotics.

Summarizing, the apparent inconsistencies in enacting experimentally devised roles are reconciled when the perceptual and the motoric behaviors are plotted on two coordinates: stability-lability of the self and rigidity-flexibility of role enactment. The latter is obviously an aspect of the number-of-roles dimension; the former is related to, if not identical with, the differentiation dimension of cognition (Werner, 1948) as applied to the self. This study supports Murphy's apt statement: "Personality is in considerable degree a matter of role behavior; even more, however, it is a matter of role perception and of self-perception in the light of the role" (1947, l. 559).

The use of self-and-role in personality scaling. Even such a prosaic, although important, topic as the prediction of college achievement is subject to analysis by means of the concepts heretofore outlined. Included in the college student's role expectations are certain actions such as going to classes, listening to lectures, writing exams, organizing abstract material, using the library, etc., and certain qualities such as friendly, cooperative, good-natured, etc.

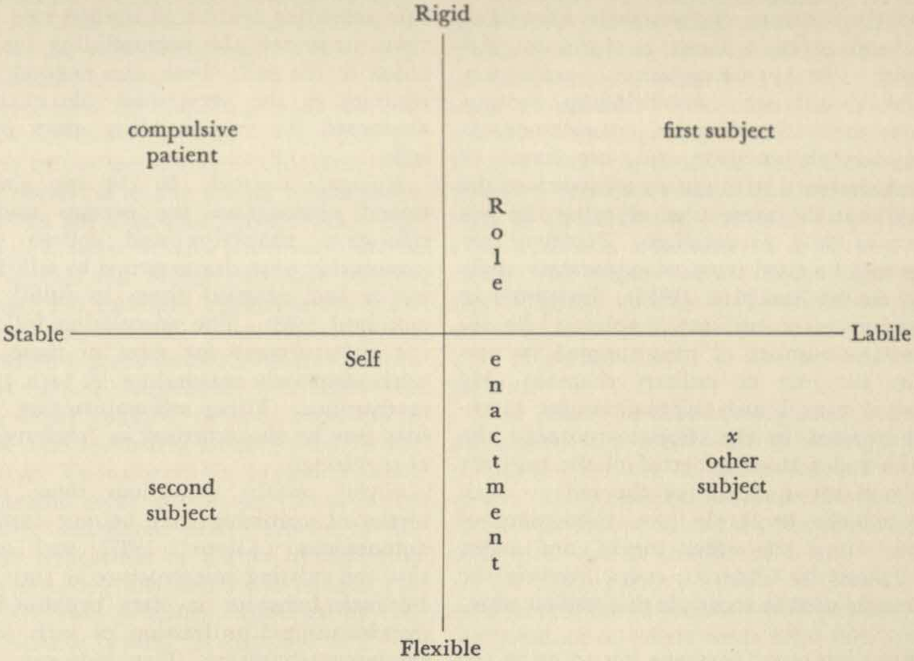


FIGURE 2.

If the actions and qualities which comprise this role are congruent with the self-concept of a particular person, then there is a high probability that he will perform according to the role expectations of the professor and other members of the college community.

Gough (1952) has devised a personality scale which gets at certain aspects of the self-concept. From the score on this scale, one can make gross predictions about the role enactment of any particular student *qua* student.

A logical analysis of the 36 items in the scale (most of which are I-statements) reveals traits of seriousness and self-confidence, consistent attitudes favoring intellectual events, and efficient scholastic habits. Since the self is a cognitive organization of traits, attitudes, and habits, it is appropriate to designate this scale as a part measure of the self. Since the expectations of the occupant of the position of student are known, we can predict the "goodness" or validity of the role enactment from this part knowledge of the self.

CONFLICT

The structure of conflict between self and incongruent role or between two incompatible roles is the ultimate source of literary productions, notably the tragedy form. In daily life such conflicts appear frequently but usually in less intense forms than those described by the novelist. We shall present here a brief social psychological analysis of the structure of conflict, utilizing the concepts introduced before.

Self-role conflict. To begin, we refer again to the principle of constancy, especially as it applies to cognition. A cognitive structure tends to maintain its organization despite forces directed toward changing it. The application of this idea to the self has already been made. The same principle applies to substructures — differentiations of the broader structure — in this case systems of role expectations. When the broader system — the qualities that make up the self — direct the organism to one set of role enactments, and the narrower system — the acts and qualities that make up the role expectations — are incongruent or disjunctive, then we have conflict. This is the quintessence of the moral dilemma.

Coincident with conflict are perturbations within the organism, phenomenally experienced as anxiety. The stable organismic state is disrupted. To re-establish constancy the person behaves in ways which make for the least change in either structure. These constancy-maintaining behaviors have been conceptualized as ego-defense mechanisms (Anna

Freud, 1937), techniques of adjustment (Cameron, 1947), defenses against anxiety (Fenichel, 1946), and psychodynamic mechanisms (Rosenzweig, 1944). Repression, projection, negativism, isolation, wish-fulfilling fantasy, displacement, rationalization, compartmentalization, overcompensation, etc., are forms of behavior designed to maintain constancy of the self while at the same time allowing the fulfillment of role expectations. Pertinent examples may be cited from an exploratory study carried out by Burchard (1953). Interested in the institution of military chaplaincy, he interviewed a number of men engaged in performing the role of military chaplain. He discovered several such self-maintaining mechanisms at work in the effort to reconcile the qualities and actions expected of the military officer and the qualities of the self — such qualities being brotherly love, abhorrence of violence, "turn the other cheek," and other traits defined by Christian codes. Among the mechanisms used to reconcile this conflict were:

Rationalization: "Someone has to carry the gospel to these boys."
Compartmentalization: "Render therefore unto Caesar the things which are Caesar's; and unto God the things that are God's."
Repression: "I don't see any conflict."
Negativism: "I'd rather not talk about it."

Another study which demonstrates the use of self-maintaining mechanisms in the enactment of an incongruent role was reported by Cousins (1951). A task involving *as if* participation in a contrived situation around a moral dilemma was presented to the subjects. He used a questionnaire-type approach in which the subject was to take the role of a student monitor pledged to enforce an unpopular and overly rigorous curfew rule. The conflict occurs when the subject has to choose between two role enactments: to report or not to report two student violators to the authorities. After reaching a decision, the subjects were instructed to select three reasons for their decisions from among a list of six reasons. The following is a generalized statement of the findings: Persons who did not fulfill the role expectations as established by the hypothetical society used self-maintaining mechanisms more frequently in explaining or justifying their conduct than did the subjects who enacted the prescribed role. That is, they were more likely to select ego-defensive items as reasons for their decision. The specific items under consideration were rationalization, displacement, and wish-fulfilling fantasy. Those who enacted the role according to the prescribed role expectations attributed the responsibility for the decision to the self; those who enacted the role contrary to the prescribed role expectations attributed the responsibility *away* from the self.

Role-role conflict. In the two above-mentioned illustrations the persons under consideration, chaplains and college students, presumably were characterized by self-structures which had enabled them to fulfill socially-approved roles. The speculation follows that the self-structures for most of these subjects were adequately maintained by such constancy mechanisms. These self-maintaining mechanisms may be characterized as "preferred modes of cognizing."

Under certain conditions these preferred modes of cognizing may become functionally autonomous (Allport, 1937) and organized into the existing self-structure as trait systems. Neurotic behavior is often explained as the overdetermined utilization of such self-maintaining mechanisms. That is to say, all roles are enacted in such a manner that the "neurotic" qualities of the role enactment stand out.

When the person has few or no self-maintaining mechanisms or if, when enacting roles incongruent with the self, the use of such mechanisms fails to re-establish constancy, then we have the preconditions for psychosis. The self is ultimately maintained, but the extreme form that the constancy mechanism takes leads to autistic, invalid, and desocialized role enactments.

The second type of social psychological conflict occurs when two substructures of the social self — two systems of role expectations — are incongruent. Warren (1949b) has written a series of propositions relating social disorganization and roles. One covariant of disorganization is the proportion of role clusters that contains incompatible roles. Of course, in a perfectly organized society, such conflicts or incompatibilities would be impossible. Since societies are not perfectly organized, institutionalized procedures are developed for reconciling or resolving conflicting role enactments. The effect of such procedures is to maintain the constancy of the social structure.

A role is a part-structure of the social self. Part-structures, like other structures, have internal organization which is subject to the constancy principle. Therefore, a role will tend to maintain itself in the face of events which threaten to disorganize it. Since the role enactment is public, the techniques by which role constancy is maintained have social sanction and are more *directly* taught and learned than is the case with the self-maintaining techniques described above.

The procedures for handling such role-role conflicts are of two broad kinds (Toby, 1952). The first is by the use of a hierarchization principle; the second is by the use of a segregation principle. In the first, the person faced with the role-role conflict utilizes the widely shared fact that certain role obligations have priority over others. For example, reporting for Selective Service has priority over reporting for work; attending grandmother's funeral has priority over school attendance, etc. However, there are other conflicts in role expectations which are not so clearly hierarchized in the social system. To maintain the preferred of two roles, a person seeks an excuse. That is, he seeks others' acceptance of his failure to discharge the obligations of a role by declaring that the competing role had a hierarchical priority. The excuse (whether supported by confirming evidence or not) is accepted as a way out of role-role conflicts and maintains the person's statuses.

Rules of etiquette make use of the hierarchization principle. Rituals are established for the orderly enactment of roles where the rights and obligations may be idiosyncratic. For example, the order of entering a doorway is prescribed by etiquette so that the woman goes first. Who is introduced to whom is likewise prescribed. Who has the right-of-way on the highway is codified. The occupant of an age-sex position, e.g., elderly lady, has priority over the occupant of another age-sex position, young man, in taking a seat on a crowded street car. These are extensions of the use of the hierachization principle.

The principle of segregation for handling role conflicts follows from the widely shared knowledge that two incompatible roles, if segregated by a time interval, may be validly enacted (Toby, 1952). It is a mechanism for maintaining role constancy. A person may occupy a position which embraces two conflicting sets of rights and obligations. The role of labor relations counselor is an example. He must serve two opposing sets of role expectations. A solution to the conflict is found by enacting the subroles *separately*. Such techniques as stalling, playing off one group against the other, repudiation of one subrole, leading a double life, etc., are cited by Toby as examples of the use of the segregation principle in resolving role conflicts.

The role-role conflict formulation is central to the study of conflicting social norms. Its applicability to the marginal man concept, for instance, is patent. An empirical method based on the theoretical system of Parsons (1949) has been contributed by Stouffer (1949) and extended by Stouffer and Toby (1951) for the study of role-role conflicts. In the experiment reported by Stouffer and Toby, a forced-choice type of questionnaire was used to get at role expectations in a number of conflicting situations. The subject had to choose between the role of friend and the role of member of the wider society. In the first instance, the orientation is particularistic, in the second, universalistic.

The results of this investigation point unmistakably to a predisposition among subjects to choose one form of role enactment over the other. The predisposition to be particularistic rather than universalistic in the choice of role enactment can be considered a trait. As such, we are dealing with a quality of the self. This underscores the thesis that valid predictions of the conduct of others must take into account aspects of ego structure as well as knowledge of competing systems of role expectations.

A METHOD FOR ASSESSING INDEPENDENT EFFECTS OF SELF AND ROLE

To the observer or behavior analyst, a person is seen as a complex entity enacting a role (Parsons and Shils, 1951). Phenomenologically, the person is a unity. If the concepts of self and role are to be justified, their experimental separation should be assayed.

It is to this task, the experimental separation of self and role, that we addressed ourselves in a large-scale study (Sarbin and Kamiya, 1952). We enlisted the services of 14 volunteers from an advanced class in psychology, 7 men and 7 women. These persons will be referred to as *social objects* in the description of the experiment. Each of these social objects enacted three roles. None of the roles was contrived, but emerged from the interaction situation. In the first role, the social object interacted with a like-sex peer (hereinafter called the auxiliary ego, after Moreno). In the second role, the social object interacted with an opposite-sex peer. In the third role, the social object interacted with an older male authority figure. The auxiliary egos were standard; that is, the *same* persons played the "other" roles *vis-à-vis* the social objects. The task was simply "to get acquainted." The interaction could be characterized as bland and nonthreatening. All the social objects were from the same psychology course, so the task given them was not unrepresentative. The persons in the interaction situations were observed and judged by college sophomores from a large lecture class in introductory psychology, $N = 526$.

Each group of subjects made judgments about one male and one female social object. The schedules were arranged so that no group of judges observed interaction situations in which the same auxiliary ego appeared. The instructions to the judges were simply to observe the interaction and to record personality judgments characteristic of the social object on a Personality Word Card, a 200-item adjective checklist. In this manner we obtained qualitative judgments of 14 persons in 3 roles. The design for males is given below. The design for the females was identical. In the interest of brevity, the descriptions of the experiment will be confined to the male objects.

Social Objects (Males)

		A	B	C	D	E	F	G	
Positions	I. *vis à vis male peer*	*x*							Roles
	II. *vis à vis female peer*	*y*		*Inter-action*					
	III. *vis à vis older male*	*z*							

Egos

In each cell are listed the 15 adjectives most commonly checked by the judging group. In cell *x*, for example, are listed the adjectives most frequently checked by the judges observing social object *A* enacting position I, male with male peer. In *y* are listed the 15 adjectives most frequently checked by an equated group of judges who observed the same social object *A* enacting position II, male with female peer. In cell *z* are listed 15 adjectives most frequently checked by an equated group of judges after observing social object *A* in position III, male with older male. Thus each cell contains the *common* adjective judgments. The horizontal dimension represents the effect of position, the vertical represents the effect of the self.

The next step was to find adjectives common to each of the social objects; by counting adjectives which appeared in the enactment of the three roles (the vertical dimension), there emerged a qualitative description of person consistency. This is, in a limited way, a description of the ego. For example, social object *B* was seen most frequently by each of the three groups of judges as quiet, sincere, calm, modest, cooperative. Social object *C* was seen as good-natured, cooperative, well-mannered, friendly, cheerful, informal, pleasant, masculine, sincere, natural.

This distillation of the columns provided us with the construct, person consistency or ego. By counting the common adjectives on the horizontal axis, we constructed a *qualitative* description of the role. A number of analyses were made, a portion of which are herein described. All the social objects are seen with some consistency in the enactment of these limited roles. There is a dispersion around a mean. Some social objects have but 5 adjectives which are common to the three roles, others as many as 13.

The role dimension provides some interesting findings. Even though verbalized requirements of the positions were identical, the qualitative role behaviors of the subjects were sufficiently different to show up in a frequency count of the rows. For at least 5 of the 7 male social objects, the following adjectives survived the subtractive analysis for each of three roles.

Role I	Role II	Role III
INFORMAL	WELL-MANNERED	AMBITIOUS
HONEST	INTELLIGENT	*frank*
SINCERE	*sociable*	cheerful
natural	*natural*	cooperative
frank	*friendly*	good-natured
sociable	cheerful	pleasant
friendly	pleasant	
cooperative	good-natured	
good-natured	cooperative	
cheerful		
pleasant		

Legend: lower case, adjectives common to three roles; italics, adjectives common to two roles; capitals, adjectives characteristic of one role.

The common adjectives for the three roles are in lower case letters. As would be expected from the similarities in the requirements of each position, the role enactments are seen as a cluster of qualities which we call "good-natured." This is one of the generalized role expectations in the middle-class student community.

The differences in the roles are interesting. When the male social objects were interacting with a like-sex peer, the qualities *informal, honest,* and *sincere* were inferred by the judges from the role enactments of the 7 social objects. When enacting the role *vis-à-vis* the opposite-sex peer, the three qualities just listed *do not* appear regularly. Instead the qualities *well-mannered* and *intelligent* survive the counting process. In the third role, where the social object appears with an older male, only one adjective meets the criterion of communality without appearing as a quality of the other roles, *ambitious*. In addition, *sociable* and *natural* are qualities common to the roles enacted *vis-à-vis peers*, and *frank* to the roles enacted *vis-à-vis males*.

In this exercise we have shown how role and ego (here defined as observed person consistency) can be empirically determined. Even though the verbalized requirements of the position were identical, certain qualities emerged from the differential role enactment, probably due to age and sex of the other.

Two further ideas emerged from this study besides those already mentioned. One is the utility of the adjective-checking method for assessing the qualitative aspects of role enactment. The second is the utility of the design which attempts to deal qualitatively with five variables simultaneously, persons (social objects), positions, roles, egos, and interactions.

FINAL STATEMENT

Persons occupy positions or statuses in interactional situations. Psychologically-considered, positions are cognitive systems of role expectations, products of learning. Role expectations are bidimensional, for every role expectation of other there is a reciprocal role expectation of self. The organized actions of the person, directed toward fulfilling these role expectations, comprise the role. Variation in role enactment is a function of at least three variables:

1. the validity of role perception (this implies the concurrent or just-prior perception and locating of the position of other and reciprocally of self);
2. skill in role enactment (related to practice in the use of *as if* behavior);
3. the current organization of the self — a cognitive structure that exercises a selective and directive effect on role perception and role enactment.

In the body of this chapter, we have tried to detail the main conceptions, wherever possible referring to empirical studies. As is usually the case with experimentation at a molar level, the rigor of the animal laboratory is absent. Nonetheless, these studies warrant inferences which give at least initial credence to predictions arising from theory. A few years ago, only impressionistic and anecdotal materials were available to document the self-role-interaction scheme of human conduct. In fact, Kluckhohn and Murray (1948, p. 268) were constrained to say: ". . . in spite of the obvious importance of role training and role playing for personality formation, this subject has barely been opened up at an empirical level." The number of empirical investigations cited in this chapter indicate that the subject has "opened up," that since 1948 social psychologists in increasing numbers have employed self, role, and interaction as key conceptions.

REFERENCES

ADORNO, T. W., *et al. The authoritarian personality.* New York: Harper, 1950.

ALBRECHT, RUTH, & SARBIN, T. R. Contributions to role-taking theory: annoyability as a function of the self. Unpublished manuscript, 1952.

ALEXANDER, G. An unexplained death coexistent with death wishes. *Psychosom. Med.*, 1943, *5*, 194-199.

ALLPORT, G. W. *Personality; a psychological interpretation.* New York: Holt, 1937.

ALLPORT, G. W. The ego in contemporary psychology. *Psychol. Rev.*, 1943, *50*, 451-479.

ARCHER, W. *Masks or faces?* New York: Longmans, Green, 1889.

ARNOLD, MAGDA B. On the mechanism of suggestion and hypnosis. *J. abnorm. soc. Psychol.*, 1946, *41*, 107-128.

ASCH, S. E. *Social psychology.* New York: Prentice-Hall, 1952.

BAKER, B. Accuracy of social perceptions of psychopathic and non-psychopathic prison inmates. Unpublished manuscript, 1954.

BALDWIN, J. M. *Social and ethical interpretations in mental development.* (3rd ed.) New York: Macmillan, 1902.

BASEDOW, H. *The Australian aboriginal.* Adelaide: Preece, 1925. (Cited in Cannon, W. B., 1942.)

BATESON, G., & MEAD, MARGARET *The Balinese character: a photographic analysis.* Special publications of the New York Academy of Sciences, Vol. II. New York: New York Academy of Sciences, 1942.

BAYLEY, NANCY Mental growth during the first three years. *Genet. Psychol. Monog.*, 1933, *14*, No. 1.

BENEDICT, RUTH Anthropology and the abnormal. *J. genet. Psychol.*, 1934, *10*, 59-82.

BENEDICT, RUTH Continuities and discontinuities in cultural conditioning. *Psychiatry*, 1938, *1*, 161-167.

BENNE, K. D., & SHEATS, P. Functional roles of group members. *J. soc. Issues*, 1948, *4*, 41-50.

BOWMAN, C. C. Role-playing and the development of insight. *Soc. Forces*, 1949, *28*, 195-199.

BROWN, J. C. An experiment in role-taking. *Amer. sociol. Rev.*, 1952, *17*, 587-597.

BROWNFAIN, J. J. Stability of the self-concept as a dimension of personality. *J. abnorm. soc. Psychol.*, 1952, *47*, 597-607.

BRUNSWIK, E. *Systematic and representative design of psychological experiments.* Berkeley: Univ. of Calif. Press, 1947.

BUGENTAL, J. F. T., & ZELEN, S. L. Investigations into the "self-concept." I. The W-A-Y technique. *J. Pers.*, 1950, *18*, 483-498.

BURCHARD, W. The role of the military chaplain. Unpublished doctoral dissertation, Univ. Calif., 1953.

CAMERON, N. The development of paranoic thinking. *Psychol. Rev.*, 1943, *50*, 219-233.

CAMERON, N. *The psychology of behavior disorders.* Boston: Houghton-Mifflin, 1947.

CAMERON, N. Role concepts in behavior pathology. *Amer. J. Sociol.*, 1950, *55*, 464-467.

CAMERON, N., & MAGARET, ANN *Behavior pathology.* Boston: Houghton-Mifflin, 1951.

CANNON, W. B. "Voodoo" death. *Amer. Anthrop.*, 1942, *44*, 169-181.

CARLSON, A. J. *The control of hunger in health and disease.* Chicago: Univ. Chicago Press, 1916.

CARTER, H. D. The development of interest in vocations. *Yearb. nat. Soc. Stud. Educ.*, 1944, *43*, 255-276.

CLINE, V. Doctoral dissertation (in preparation), Univ. Calif., Berkeley, 1953.

COTTRELL, L. S., JR. The adjustment of the individual to his age and sex roles. *Amer. sociol. Rev.*, 1942, *7*, 618-625.

COTTRELL, L. S., JR. Some neglected problems in social psychology. *Amer. sociol. Rev.*, 1950, *15*, 705-712.

COTTRELL, L. S., JR., & DYMOND, ROSALIND F. The empathic responses: a neglected field for research. *Psychiatry*, 1949, *12*, 355-359.

COTTRELL, L. S., JR., & GALLAGHER, RUTH *Developments in social psychology, 1930-1940.* New York: Beacon House, 1941.

COUSINS, A. N. Social equilibrium and the psychodynamic mechanisms. *Soc. Forces*, 1951, *30*, 202-209.

COUTU, W. Role-playing vs. role-taking: an appeal for clarification. *Amer. sociol. Rev.*, 1951, *16*, 180-187.

CRAWLEY, E. *The mystic rose.* (Rev. by T. Besterman.) London: Methuen, 1902.

DEVEREUX, G. Social and cultural implications of incest among the Mohave. *Psychoanal. Quart.*, 1939, *8*, 510-533.

DUBOIS, CORA Attitudes toward food and hunger in Alor. In L. Spier and others, *Language, culture, and personality: essays in memory of Edward Sapir.* Menasha: Sapir Memorial Publ. Fund, 1941.

DYMOND, ROSALIND F. A scale for the measurement of empathic ability. *J. consult. Psychol.*, 1949, *13*, 127-133.

FENICHEL, O. *Psychoanalytic theory of neurosis.* New York: Norton, 1946.

FORD, C. S. Role of a Fijian chief. *Amer. sociol. Rev.*, 1938, *3*, 541-550.

FREDERICSON, E. Competition: the effects of infantile experience upon adult behavior. *J. abnorm. soc. Psychol.*, 1951, *46*, 406-410.

FREUD, ANNA *The ego and the mechanisms of defense.* London: Hogarth Press, 1937.

FREUD, S. *The ego and the id.* London: Hogarth Press, 1935.

GOLDMAN, F. Breastfeeding and character formation. II. The etiology of the oral character in psychoanalytic theory. *J. Pers.*, 1950, *19*, 189-197.

GOUGH, H. G. A sociological theory of psychotherapy. *Amer. J. Sociol.*, 1948, *53*, 359-366.

GOUGH, H. G. Predicting success in graduate training: a progress report. Berkeley: Univ. Calif. Instit. Personality Assessment and Res., 1950 (mimeographed).

GOUGH, H. G. The construction of a personality scale to predict academic achievement in introductory psychology courses. *Amer. Psychologist*, 1952, *7*, 367.

GOUGH, H. G., & PETERSON, D. R. The identification and measurement of predispositional factors in crime and delinquency. *J. consult. Psychol.*, 1952, *16*, 207-212.

HACKER, HELEN M. Women as a minority group. *Soc. Forces*, 1951, *30*, 60-69.

HARSH, C. M. An inventory study of categories of annoyance. Unpublished doctoral dissertation, Univ. Calif., Berkeley, 1936.

HARTLEY, E. L., & HARTLEY, R. E. *Fundamentals of social psychology.* New York: Knopf, 1952.

HARTLEY, E. L., & KRUGMAN, DOROTHY C. Note on children's social role perception. *J. Psychol.*, 1948, *26*, 399-405.

HARTLEY, E. L., ROSENBAUM, M., & SCHWARTZ, S. Children's use of ethnic frames of reference: an exploratory study of children's conceptualizations of multiple ethnic membership. *J. Psychol.*, 1948a, *26*, 367-386.

HARTLEY, E. L., ROSENBAUM, M., & SCHWARTZ, S. Children's perceptions of ethnic group membership. *J. Psychol.*, 1948b, *26*, 387-398.

HASTORF, A. H., & BENDER, I. E. A caution respecting the measurement of empathic ability. *J. abnorm. soc. Psychol.*, 1952, *47*, 564-576.

HERSKOVITS, M. J. *Man and his works.* New York: Knopf, 1948.

HILGARD, E. R. Human motives and the concept of the self. *Amer. Psychologist,* 1949, *4,* 374-382.

HUGHES, E. C. Personality types and the division of labor. *Amer. J. Sociol.,* 1928, *33,* 754-768.

HUGHES, E. C. Institutional office and the person. *Amer. J. Sociol.,* 1937, *43,* 404-413.

HUNT, J. McV. The effects of infant feeding-frustration upon adult hoarding in the albino rat. *J. abnorm. soc. Psychol.,* 1941, *36,* 338-360.

HUNT, J. McV., SCHLOSBERG, H., SOLOMON, R. L., & STELLAR, E. Studies of the effects of infantile experience on adult behavior in rats. I. Effects of infantile feeding frustration on adult hoarding. *J. comp. physiol. Psychol.,* 1947, *40,* 291-304.

JACOBSON, E. *Progressive relaxation.* (Rev. Ed.) Chicago: Univ. Chicago Press, 1938.

JAMES, W. *The principles of psychology.* New York: Henry Holt, 1890.

KOFFKA, K. *Principles of Gestalt psychology.* New York: Harcourt, Brace, 1935.

KOMAROVSKY, MIRA Cultural contradictions and sex roles. *Amer. J Sociol.,* 1946, *52,* 184-189.

KRECH, D., & CRUTCHFIELD, R. S. *Theory and problems of social psychology.* New York: McGraw-Hill, 1948.

LECKY, P. *Self-consistency: a theory of personality.* New York: Island Press, 1945.

LEWIN, K. *A dynamic theory of personality; selected papers.* Transl. by D. K. Adams and K. E. Zener. New York: McGraw-Hill, 1935.

LEWIS, J. H., & SARBIN, T. R. Studies in psychosomatics: the influence of hypnotic stimulation on gastric hunger contractions. *Psychosom. Med.,* 1943, *5,* 125-131.

LINDESMITH, A. R., & STRAUSS, A. L. *Social psychology.* New York: Dryden Press, 1949.

LINTON, R. *The cultural background of personality.* London: Routledge and Kegan Paul, 1947.

McCLELLAND, D. C. *Personality.* New York: Wm. Sloane Associates, 1951.

MATHEWS, RAVENNA, HARDYCK, C., & SARBIN, T. R. Contributions to role-taking theory. VI. Self-organization as a factor in the performance of selected cognitive tasks. *J. abnorm. soc. Psychol.,* 1953, *48,* 500-503.

MAZELIS, SARAH The relation of role-diffusion to work productivity. Unpublished manuscript, 1953.

MEAD, G. H. *Mind, self, and society.* Chicago: Univ. Chicago Press, 1934.

MERTON, R. K. Bureaucratic structure and personality. *Soc. Forces,* 1940, *18,* 560-568.

MILLER, N. E., & DOLLARD, J. *Social learning and imitation.* New Haven: Yale University Press, 1941.

MORENO, J. L. *Who shall survive?* Nervous and mental disease monograph series, No. 58. Washington, D. C.: Nervous and Mental Disease Publ. Co., 1934.

MORENO, J. L. *Psychodrama.* Vol. I. New York: Beacon House, 1946.

MOTZ, ANNABELLE B. The role conception inventory. *Amer. sociol. Rev.,* 1952, *17,* 465-471.

MURPHY, G. *Personality: a biosocial approach to origins and structure.* New York: Harpers, 1947.

NABOISEK, H. Validation of a method for predicting role expectations in group therapy. Unpublished doctoral dissertation, Univ. Calif., Berkeley, 1953.

NEIMAN, L. J., & HUGHES, J. W. The problem of the concept of role — a re-survey of the literature. *Soc. Forces,* 1951, *30,* 141-149.

NEWCOMB, T. M. *Social psychology.* New York: Dryden Press, 1950a.

NEWCOMB, T. M. Role behaviors in the study of individual personality and of groups. *J. Pers.,* 1950b, *18,* 273-289.

NOTCUTT, B., & SILVA, A. L. M. Knowledge of other people. *J. abnorm. soc. Psychol.,* 1951, *46,* 30-37.

NOTTINGHAM, ELIZABETH K. Toward an analysis of the effects of two world wars on the role and status of middle-class women in the English-speaking world. *Amer. sociol. Rev.,* 1947, *12,* 666-675.

ORT, R. S. A study of role-conflicts as related to class level. *J. abnorm. soc. Psychol.,* 1952, *47,* 425-432.

PARK, R. E., & BURGESS, E. W. *Introduction to the science of society.* (2nd Ed.) Chicago: Univ. Chicago Press, 1924.

PARSONS, T. *Essays in sociological theory.* Glencoe, Ill.: Free Press, 1949.

PARSONS, T. *The social system.* Glencoe, Ill.: Free Press, 1951.

PARSONS, T., & SHILS, E. A. *Toward a theory of action.* Cambridge: Harvard Univ. Press, 1951.

PIAGET, J. *The language and thought of the child.* New York: Harcourt, Brace, 1926.

RIVERS, W. H. R. *Medicine, magic, and religion.* New York: Harcourt, Brace, 1924.

ROGERS, C. R. Some observations on the organization of personality. *Amer. Psychologist,* 1947, *2,* 358-368.

ROKEACH, M. Generalized mental rigidity as a factor in ethnocentrism. *J. abnorm. soc. Psychol.,* 1948, *43,* 259-278.

ROSE, A. M. The adequacy of women's expectations for adult roles. *Soc. Forces,* 1951, *30,* 69-77.

ROSENZWEIG, S. An outline of frustration theory. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald Press, 1944. Vol. I, pp. 379-388.

ROSENZWEIG, S., & SARASON, S. An experimental study of the triadic hypothesis: reaction to frustration, ego-defense, and hypnotizability. *Charact. Pers.,* 1942, *11,* 1-19.

SARBIN, T. R. The concept of role-taking. *Sociometry,* 1943, *6,* 273-285.

SARBIN, T. R. Contributions to role-taking theory. I. Hypnotic behavior. *Psychol. Rev.,* 1950, *57,* 255-270.

Sarbin, T. R. Contributions to role-taking theory. III. A preface to a psychological analysis of the self. *Psychol. Rev.,* 1952a, *59,* 11-22.

Sarbin, T. R. Contributions to role-taking theory. Role-expectations as qualities. Unpublished manuscript, 1952b.

Sarbin, T. R., & Farberow, N. L. Contributions to role-taking theory. II. A clinical study of self and role. *J. abnorm. soc. Psychol.,* 1952, *47,* 117-125.

Sarbin, T. R., & Hardyck, C. Contributions to role-taking theory: role-perception on the basis of postural cues. Unpublished manuscript, 1953.

Sarbin, T. R., & Kamiya, J. Contributions to role-taking theory: the experimental separation of qualities of self and of role. Unpublished manuscript, 1952.

Sarbin, T. R., & Rosenberg, B. G. Contributions to role-taking theory. IV. A method for a qualitative analysis of the self. *J. soc. Psychol.,* 1953, in press.

Sarbin, T. R., & Stephenson, R. W. Contributions to role-taking theory: authoritarian attitudes and role-taking skill. Unpublished manuscript, 1952.

Sarbin, T. R., & Van Spanckeren, W. J., Jr. Contributions to role-taking theory: the relation of number of remembered parent-figures to phenomenal self descriptions. Unpublished manuscript, 1953.

Sarbin, T. R., & Williams, J. D. Contributions to role-taking theory. V. Role perception on the basis of limited auditory stimuli. Unpublished manuscript, 1953.

Sargent, S. S. *Social psychology.* New York: Ronald Press, 1950.

Schilder, P. F. *The image and appearance of the human body.* London: Paul, Trench, Trubner, 1935.

Schuessler, K. F., & Cressey, D. F. Personality characteristics of criminals. *Amer. J. Sociol.,* 1950, *55,* 476-484.

Schultz, J. H. *Das Autogene Training* (Konzentrative Selbstentspannung). Leipzig: Thieme, 1932.

Sears, R. R. Ordinal position in the family as a psychological variable. *Amer. sociol. Rev.,* 1950, *15,* 398-401.

Shirley, M. M. *The first two years: a study of 25 babies.* Vol. II. *Intellectual development.* Minneapolis: Univ. Minn. Press, 1933.

Sickles, W. R., & Hartmann, G. W. The theory of order. *Psychol. Rev.,* 1942, *49,* 403-421.

Smith, M. B. The phenomenological approach in personality theory: some critical remarks. *J. abnorm. soc. Psychol.,* 1950, *45,* 516-522.

Stellar, E., & Morgan, C. T. The roles of experience and deprivation in the onset of hoarding behavior in the rat. *J. comp. Psychol.,* 1943, *36,* 47-55.

Stouffer, S. A. An analysis of conflicting social norms. *Amer. sociol. Rev.,* 1949, *14,* 707-717.

Stouffer, S. A., & Toby, J. Role conflict and personality. *Amer. J. Sociol.,* 1951, *56,* 395-406.

Sullivan, P. L. An investigation of the conceptual properties governing the categorization of people, with special reference to certain social attitudes and values. Unpublished doctoral dissertation, Univ. Calif., Berkeley, 1950.

Symonds, P. M. *The ego and the self.* New York: Appleton-Century-Crofts, 1951.

Toby, J. Some variables in role-conflict analysis. *Soc. Forces,* 1952, *30,* 324-327.

Tolman, E. C. Prediction of vicarious trial and error by means of the schematic sowbug. *Psychol. Rev.,* 1939, *46,* 318-336.

Tolman, E. C. A theoretical analysis of the relations between sociology and psychology. *J. abnorm. soc. Psychol.,* 1952, *47,* 291-299.

Vaihinger, H. *The philosophy of "as-if."* London: Paul, Trench, Trubner, 1924.

Waller, W. W. *The sociology of teaching.* New York: Wiley, 1932.

Wallin, P. Cultural contradictions and sex roles: a repeat study. *Amer. sociol. Rev.,* 1950, *15,* 288-293.

Warner, W. L. *A black civilization: a social study of an Australian tribe.* New York: Harpers, 1937.

Warren, R. L. Cultural, personal, and situational roles. *Sociol. soc. Res.,* 1949a, *34,* 104-111.

Warren, R. L. Social disorganization and the interrelationship of cultural roles. *Amer. sociol. Rev.,* 1949b, *14,* 83-87.

Werner, H. *Comparative psychology of mental development.* (Rev. ed.) New York: Follett, 1948.

Young, K. *Social psychology.* New York: Crofts, 1945.

Young, K. *Personality and problems of adjustment.* (Rev. ed.) New York: Appleton-Century-Crofts, 1952.

Part 3

Research Methods

CHAPTER 7

Experiments: Their Planning and Execution

Allen L. Edwards
The University of Washington

Psychologists do research because they seek answers to problems or because they wish to test answers that have been suggested. For example, suppose we pose this problem: Is it better to present one or both sides of a controversial issue, if we desire to shift attitudes in one direction only? A search through the available literature may reveal that (1) the question has not been considered before; that is, no answers have been suggested; (2) the question has been considered, but more than one answer has been suggested; (3) the question has been considered and there is agreement among authorities as to the answer.

In case a problem has not been previously considered, research might be undertaken to determine an answer to the problem. If a problem has been considered, but various answers have been suggested, then research might be undertaken to determine which of the various answers is most adequate under specified conditions.

Even when authorities are agreed upon a single answer to a problem this does not necessarily mean that the answer is correct. Authorities are not infallible. How did they arrive at the answer they offer? If the answer is based upon speculation and theory it must be granted that these do not provide as sound a basis for an answer as might be desired.

Speculation is valuable and may provide good hunches, but hunches are still hunches and must be regarded as tentative answers. Theory is also valuable in raising hypotheses or tentative answers to problems. Tentative answers can become definite answers, however, only when tested by empirical research.

CLASSIFICATION OF PSYCHOLOGICAL RESEARCH

Research is done to seek an answer to a question or to a problem, or to test answers which have been suggested. A convenient classification of psychological research has been provided by Edwards and Cronbach (1952) in terms of the kinds of problems which psychologists raise. These authors propose that psychological research may be classified as (1) survey research, (2) technique research, (3) applied research, and (4) critical research.

SURVEY RESEARCH

Survey research is concerned with the rather general problem of what variables will be relevant to or relate to another given variable. Survey research is exploratory in nature and is most useful in the preliminary stages of investigation of a particular problem. This is the time at which the psychologist is relatively ignorant of the variables that may be of significance.

Suppose a psychologist is interested in determining what variables are relevant to opinions held on a given issue. He does not know, for example, whether the reasons people give for holding the opinions they hold will be relevant or not. The same reasons may be given by those holding opposed opinions. If this is true, then the psychologist cannot very well say that the reasons given are relevant to the opinions held. On the other hand, if it is found that reasons given are in some way related to the opinions held, then this variable may be considered relevant to the problem.

Similarly, survey research might investigate whether political party affiliation, socioeconomic status, educational level, sex, intellectual level, personality differences, and so forth, are in any way related to opinions held on a defined issue. Numerous examples of this kind of survey research may be found in Cantril (1944).

The hunches or tentative answers guiding the psychologist in his selection of the variables to be investigated in survey research are unusually vague. The variables observed may be determined to a large extent by the hope that if a sufficiently large number are investigated, some of them may prove to be relevant.

TECHNIQUE RESEARCH

Technique research is concerned with the problem of the method or technique to be used in the making of observations. It may be desired, for example, to do research upon attitudes on some issue. The question arises as to the method by which attitudes are to be studied. If techniques or methods are not available for making observations of this variable, they must be developed before the research can go ahead. Examples of technique research may be found in the work of Thurstone and Chave (1929), Guttman (1947), Likert (1932), and Edwards and Kilpatrick (1948). These studies are all concerned with the development of methods of measuring social attitudes.

When several methods or techniques of observing a particular variable are available, technique research may be concerned with a comparison of these methods. Technique research, in this instance, may provide answers to such questions as: Which method has the greatest reliability? Which method is the simplest to use? To what extent do the various methods yield comparable results? The work of Edwards and Kenney (1946) involving a comparison of the Likert (1932) and Thurstone and Chave (1929) methods of attitude scale construction is an example of this kind of technique research. The comparisons of different psychological scaling methods by Hevner (1930), Saffir (1937), and Attneave (1949) are other examples of technique research.

APPLIED RESEARCH

Applied research is concerned with problems of a practical administrative nature. Suppose we are confronted with the problem of choosing among several methods for changing attitudes in a desired direction. Which method will best attain the objective? Research may provide us with an answer and thus guide our choice.

Several methods may be proposed for the selection of leaders in a group situation. Which of these methods gives the best results? Which method will select the largest number of good leaders and the smallest number of poor leaders? If research shows that the several methods are equally effective, then administrative action may be taken in terms of the relative costs of the methods, or in terms of the simplicity of the methods.

CRITICAL RESEARCH

Critical research is research designed to test a hypothesis stemming from theory. A certain theory suggests that the outcome of the research should be in a certain direction. We have, in other words, from prior considerations, some belief that if the research is done, it should yield a particular answer. The research is then done in order to determine whether the outcome is compatible with theoretical expectations.

OVERLAPPING OF CLASSES

Although psychological research may be classified as survey, technique, applied, and critical research, it should be recognized that there may be a great deal of overlapping among the classes. Survey research may have practical or applied aspects, once it is completed. Examination of the results of survey research may provide a basis for administrative decisions, although this was not contemplated when the research was planned. Technique research may also have applied aspects, and applied research may have theoretical as well as practical implications.

If we ask the question, however, as to why a particular piece of research is to be done, prior to the actual research itself, we may then find that the motivation for doing the research enables us to place it in one or another of the four classes described. It is being done to discover relevant variables, or to investigate a method of making observations, or to determine what action should be taken, or to test some hypothesis suggested by theoretical considerations.

Experiments and Experimentation

Experiments are particular ways of making observations. An experiment differs from other research methods in that the experimenter has some degree of control over the variables involved and the conditions under which the variables are observed.

To determine whether socioeconomic status is related to opinion on a given issue does not involve experimentation, but the answer can be determined by survey research or the making of pertinent observations. In this problem, the investigator is studying two variables, socioeconomic status and opinions held, as they exist in the group under observation. The investigator has no control over the socioeconomic status of the subjects nor of their opinions. Socioeconomic status is already fixed for each subject; it cannot be manipulated or changed directly by the research worker. Similarly with respect to opinions. The opinions that are held by the members of the group are the opinions that are held. The investigator can only determine what these opinions are and then see whether they are related to socioeconomic status.

In an experiment, however, the experimenter has control over certain variables. These variables are called the *independent* variables. Independent variables are variables which the experimenter himself manipulates or changes. As the independent variables are changed, the experimenter observes other variables to see whether they are related to the changes introduced. These variables are called the *dependent* variables. The problem of the experiment is to determine the relationship between the independent and dependent variables.

An experiment might be conducted to determine whether a certain skill, as measured by a psychological test, can be most effectively taught by means of a given film strip or by means of a specified lecture. The film strip or lecture would be the independent variable and the performance of the subjects, as measured by the test, would be the dependent variable. The experiment is designed to determine if there is any relationship between the dependent variable (performance) and the independent variable (film strip versus lecture). The control exercised by the experimenter has to do with the independent variable. He can determine the conditions under which the independent variable is to be introduced and he can control other factors that might influence the dependent variable.

The advantages of making observations under controlled conditions over observations made without such control have been pointed out by Woodworth (1938, p. 2):

> 1. The experimenter makes the event happen at a certain time and place and so is fully *prepared* to make an accurate observation.
>
> 2. Controlled conditions being *known* conditions, the experimenter can set up his experiment a second time and repeat the observation; and, what is very important in view of the social nature of scientific investigation, he can report his conditions so that another experimenter can duplicate them and check the data.
>
> 3. The experimenter can systematically *vary* the conditions and note the concomitant variation in the results.

It was customary in the earlier days of experimentation to believe that the experimenter must hold all variables (independent) constant except one. The "rule of the single variable" was formulated before more efficient statistical methods and experimental designs had been developed. In modern experiments, with the statistical tools now available, it is possible to handle several independent variables in the same design and also to have as many dependent variables as seem necessary. Experimental designs of this sort will be discussed later.

TESTS OF SIGNIFICANCE

Since experiments are one of the methods of research, it should be clear that experiments are designed to yield answers to problems. The statement of objectives of an experiment is a statement of the question or problems which the experiment is planned to answer.

It is sometimes said that experiments are planned in order to test hypotheses. This is true in the broad sense that a hypothesis is a statement of the possible nature of the relationship between two or more variables. It is a tentative answer to the question of what the relationship is. Now, if a hypothesis is regarded as a statement of the relationship between two or more variables, and if variables are defined as the things which the research worker observes, then it is perfectly obvious that the test of a hypothesis must involve the making of observations. In other words, we make observations of each of the variables involved in the hypothesis and then determine whether the relationship is of the nature stated by the hypothesis.

The hypotheses or tentative answers concerning relationships which the experimenter has prior to the making of observations under the experimental conditions can be expressed more precisely in statistical form, once the observations have been made. Outcomes of experiments are usually expressed in terms of means, standard deviations, proportions, correlation coefficients, and other statistics. These statistics summarize the results of the experiment in a convenient form.

A test of significance can then be made in which the statistic expressing the results of the experiment is compared with a theoretical distribution of results to be expected under a specified hypothesis. This hypothesis is often called a *null* hypothesis (Fisher, 1942). Theoretical distributions frequently used in tests of significance include the t distribution, the F distribution, and the χ^2 distribution. On the basis of the observations made during the course of the experiment, an observed value of t, F, or χ^2 may be computed. This observed value is then compared with the available tables (Fisher and Yates, 1949) of these values to determine the probability of obtaining the observed value under the hypothesis being tested.

The test of significance thus results in a probability value. When the probability given by the test of significance is small, say .05 or less, then we say that the obtained value is significant. When the probability given by the test of significance is large, say greater than .05, we say that the obtained value of *t*, *F*, or χ^2 is not significant. The important point is that we not only need to know the outcome of the experiment, expressed in some statistic such as a mean, correlation coefficient, standard deviation, or variance, but that we often find it desirable to evaluate the outcome in terms of some test of significance.

ONE- AND TWO-TAILED TESTS OF SIGNIFICANCE

The nature of the test of significance will depend upon the hypothesis formulated by the experimenter prior to the actual experiment. We may distinguish between two major cases. If the experimenter's hypothesis concerns the *direction* of the relationship to be found, the results will be evaluated in terms of a one-tailed test of significance. The experimenter, in this instance, is interested only in determining whether the results are in the direction specified by the hypothesis. On the other hand, the experimenter may be interested in determining whether *any* relationship exists, regardless of its direction. In this instance, the outcome of the experiment would be evaluated by means of a two-tailed test of significance.

One- and two-tailed tests of significance are discussed in terms of psychological experiments by Jones (1952) and Marks (1951). The terms one- and two-tailed tests refer to whether one or both tails of the theoretical distributions, discussed previously, are to be used in arriving at the probability value.

Consider the case, for example, where the results of an experiment are expressed in terms of the correlation coefficient between a dependent and an independent variable. The correlation coefficient is a measure of the relationship between two variables when they are linearly related. It can range in value from 1.00, indicating a perfect positive relationship, through zero, indicating no relationship, to -1.00, indicating a perfect negative relationship. Suppose the observed value of the correlation coefficient as determined by the observations made in the experiment is .48.

If the experimenter had started out with the hypothesis that the relationship between the two variables should be *positive*, the observed value of the correlation coefficient would be evaluated by means of a one-tailed test of significance. The experimenter, in this case, is interested in determining whether he may conclude that the relationship is significantly greater than zero in a positive *direction*. On the other hand, if the experimenter merely wanted to determine whether *any* relationship existed between the two variables, i.e., whether a positive or a negative relationship existed, the observed value of the correlation coefficient would be evaluated by means of a two-tailed test of significance. In essence, the experimenter is interested in determining whether the correlation coefficient differs significantly from zero in either a positive or a negative direction.

Consider another example, in which one group of subjects is tested under experimental condition *A* and another group of subjects is tested under experimental condition *B*. The independent variable here is the experimental condition, and the dependent variable is a measure of performance of the subjects under the experimental conditions. The outcome of the experiment is summarized in terms of the means of the groups on the dependent variable along with a measure of the variation in performance, i.e., the standard deviation or its square, the variance.

If the experimenter, prior to the experiment, had formulated a hypothesis that the mean of, say, the *A* group should be higher than the mean of the *B* group, then the difference between the means would be evaluated by means of a one-tailed test of significance. In essence, the experimenter is stating that the average performance should be positively related to the *A* condition. He is interested in determining whether the mean of the *A* group is significantly greater than the mean of the *B* group. On the other hand, if the experimenter had no prior hypothesis as to which group should be superior, the difference between the means would be evaluated in terms of a two-tailed test of significance. In essence, the experimenter is interested in determining whether the relationship is such that either the *A* group will be superior to the *B* group, or vice versa.

TYPE I AND TYPE II ERRORS

The hypothesis tested may be true in fact, but the test of significance may erroneously result in the rejection of the hypothesis. For example, the experimenter may be testing the null hypothesis that two means do not differ significantly, and he is prepared to reject this hypothesis on the basis of a two-tailed test of significance. If it should actually be true that the two means do not differ, but the test of significance results in the rejection of the null hypothesis, this is called a type I error. On the other hand, the two means may actually differ, but the test of significance may fail to result in the rejection of the null hypothesis. Acceptance of the null hypothesis under this condition is called a type II error.

It is obviously desirable that any given test of significance result in the rejection of very few true hypotheses and, at the same time, result in the rejection of as many false hypotheses as possible. The probability of rejection of a true hypothesis can be set by the experimenter. For example, if the experimenter decides to view a probability value of .05 as a basis for rejecting the hypothesis tested, then the frequency of type I errors will be 5 in 100. If the experimenter demands a probability of .01 before rejecting the hypothesis, then the frequency of type I errors will be 1 in 100. It should be clear that if the experimenter demands an extremely small probability before rejecting a hypothesis, the result will be a decrease in the type I errors he makes. Yet, at the same time, he will increase the frequency of type II errors he makes. The probability values that are commonly used in research work are those of .05 and .01.

Type I and type II errors are discussed in greater detail from the point of view of statistical considerations by Hoel (1947), Mood (1950), and Johnson (1949). Less technical discussions can be found in Walker (1943), Tippett (1941), Cochran and Cox (1950), and Marks (1951). See also Operating Characteristics or Power Curves in Chapter 8.

VARIABLES IN PSYCHOLOGICAL RESEARCH

The variables which psychologists observe in research may be represented as response variables, stimulus variables, and organismic variables. These three terms are used in a broad sense. Response, for example, is used to designate any kind of behavior or action of the organism. Some responses may be very simple in nature, such as a finger flexion, and others may be highly complex, such as the behavior we call aggression. Stimulus variables also range in complexity. In some experiments the stimulus may be of a relatively simple kind, such as a flash of light. In other experiments the stimulus may represent a complex social situation.

Any property or characteristic of the organism may, in a given experiment, be called an organismic variable. A clear-cut example of an organismic variable would be height or weight, neither of which seems properly to fit within the class of stimulus or response variables. Other organismic variables may be designated upon the basis of prior experience of the individuals. Subjects in an experiment, for example, may be classified into groups in terms of their degree of prior education. Sometimes an experiment may incorporate into its design an organismic variable based upon some previous observation of behavior of the subjects. For example, it may be useful in some experiments to classify subjects upon the basis of their present attitudes toward some issue, or on personality traits. This classification, in turn, might be based upon prior observation of a response variable, i.e., scores upon an attitude scale or a personality inventory. In all such cases as those described, the organismic variable is an independent variable. It is one of the factors introduced into the design by the selection or classification of subjects by the experimenter.

In terms of these three kinds of variables, we may attempt a classification of the specific problems of research which the psychologist undertakes. Introductory texts in general psychology usually state the principle that

$$R = f\,(O, S),$$

which means that response or behavior is believed to be a function both of properties of the organism and of the stimulus situation.

But not all research is concerned with the formulation $R = f\,(O, S)$. Technique research, as described earlier, is often concerned with a single variable and the problem of recording or measuring that variable more precisely. It is conceivable that other psychological research problems may be concerned with any two or more of the three kinds of variables considered in all possible permutations (Edwards, 1950b). As a single illustrative example, consider the case where the research interest is in relating socioeconomic status to educational level. This might be considered as an instance of an *O-O* research problem. That which differentiates the social psychologist from other psychologists is his interest in *O*, *S*, and *R* variables that are social in nature.

One of the gravest tasks which confronts the experimenter is the defining of the specific variables involved in a problem. For research purposes, variables must be defined in terms of observations or operations upon observations.

Research and experimentation consist of attempting to relate two or more such sets of observations (variables). Consequently, variables must be defined in terms of the observations made and related. A variable can mean no more than the set of observations or operations upon the observations which define it. Hypotheses, it may be recalled, are tentative answers concerning the relationships between variables which are to be tested by research and experimentation.

QUANTIFICATION OF VARIABLES

The problem confronting the experimenter in the primary stages of planning his work is that of determining how and in what way the variables he intends to work with are to be observed and recorded. It is the aim of all branches of science to seek quantification of observations. Quantification of observations is a step toward quantification of the relationships between variables. As the formulation $R = f(O, S)$ indicates, it is desired to express behavior as some mathematical function of various organismic and stimulus variables.

With some variables, the problem of quantification is relatively simple. Various techniques have been devised and constructed for the measurement of variables with which the psychologist deals. Suppose that in an experiment the stimulus variable consists of a set of weights which are to be lifted and judged in comparison with some standard weight. The weights themselves can be measured on a physical continuum, i.e., the quantification of the stimulus variable is a simple matter of determining the actual weight of each stimulus. But the problem of quantification of the response variable, i.e., the judgment made by the subject, is much more complex. Suppose the standard stimulus has been assigned a middle value in the series of weights and the subjects are asked to rate the degree of weight of the comparison stimuli in terms of a nine-point scale extending on both sides of the standard. No physical continuum exists or is known for the series of judgments or responses obtained from the subjects.

In this instance, however, quantification can be obtained on a psychological continuum by means of various techniques of psychological scaling which have been developed (Guilford, 1938; Attneave, 1949; Edwards and Thurstone, 1952. See also Chapter 9). Psychological scaling methods have been given a great deal of attention by research workers and it is now possible to scale or measure judgments with a rather high degree of accuracy and reliability. These methods are of particular value when physical measurement of a variable is not possible.

Psychological tests represent another valuable source of quantification for response and organismic variables of various kinds. A vast number of psychological tests exist for the purpose of quantifying skills, abilities, personality attributes, and other variables. Psychological tests of this nature are described by Goodenough (1949) and Cronbach (1949). The methodology of test construction, as described by Gulliksen (1950), has been highly developed and can be applied to research problems for the construction of new devices or instruments for quantifying behavior where none already exist.

Rating methods, described by Guilford (1936), are another valuable set of devices for the research worker. These methods may be used for a vast number of problems in which it is desired to obtain a numerical value for some complex behavior which can be judged by observers but not otherwise measured. Ranking methods may also be used for this purpose. Kendall (1948) has described statistical methods appropriate for the analysis of ranked observations. For further discussion see Nonparametric Methods in Chapter 8.

As important as the defining of variables is the problem of the reliability of the observations themselves. By reliability is meant the extent to which we would obtain comparable observations in the same circumstances. If two equally trained observers cannot rate the behavior of subjects in an experiment in such a way that their ratings agree, we say that the rating system is unreliable or has a low degree of reliability. Obviously, with unreliable observations, we cannot hope to have much faith in any relationships established for this set of observations with some other variable. Experimentation and research demand that the observations made be reliable.

Physical measurements possess a high degree of reliability, and psychological tests can usually be so constructed that the scores obtained from them are reliable. Various techniques have been devised for increasing the reliability of ratings (Guilford, 1936).

SELECTION OF SUBJECTS

In discussing selection of subjects to be used in an experiment, we shall consider two cases: one where the experimental design involves an organismic variable as an independent variable and the other where it does not.

If we took a group of unselected college students and measured them with respect to almost any variable, we would find that they showed considerable individual differences or variation. They would differ with respect to height, weight, abilities, attitudes, skills, previous experience, interests, socioeconomic status, traits, habits, and so on. A measure of this variation or of individual differences would be given by the standard deviation or by the square of the standard deviation, called the variance, of the distribution of observations. The arithmetic mean of the distribution would give a characteristic value which we assume represents the typical member of the group.

Now if we took another group of unselected subjects from the population at large, we could also find the standard deviation of their distribution of observations which would be a measure of their variation. And we could find the arithmetic mean of the distribution, which we again may assume would represent the typical member of this unselected group. In comparing the means and standard deviations obtained from our two groups, we should probably arrive at two conclusions, regardless of the nature of the particular observations made. The groups would differ with respect to their arithmetic means, but the direction of this difference would not necessarily be systematic. We might expect the college group to have a higher mean than the unselected group from the general population with respect to abilities, intellectual level, socioeconomic status, and on any other variable that is positively related to educational status. But on some of the variables which may be negatively related to educational status, we would find that the college group has a lower mean than the group from the general population.

With respect to the standard deviations, however, it would be safe to predict that, in general, the college group would show less variation than the unselected group from the general population. For example, the college group would vary less in height and in age than the unselected group, and this would be true for almost any other variable that we might name.

Now, when an experiment is done with a sample or group of subjects and certain relationships are found between the variables investigated, it is usually the desire of the experimenter to be able to generalize his results beyond the particular group which has been studied. If a certain result is found when male college students are subjects, to what extent may we expect this result to hold true with female subjects? With high-school subjects? With older subjects? Strictly speaking, it is possible to generalize beyond a sample only to other groups that are comparable to the group investigated. And comparable, in this instance, means with respect to the organismic variables that characterize the group investigated.

If the group investigated consists of subjects who are above average or below average or all of approximately average level with respect to some organismic variable, then we cannot be sure, without further investigation, that the results obtained will hold for other groups of different levels of the organismic variable. The organismic variables may be important independent variables in the sense that they are related to or influence the results observed with respect to the dependent variables.

The problem here is similar to that found in experimental work in other branches of science. A well-planned experiment in agriculture, for example, may show that a specified amount of fertilizer applied to units of land in a section of Iowa results in greater production of corn than any other specified amount of the fertilizer tested in the experiment. The varying amounts of fertilizer, in this experiment, correspond to the independent variable, and the amount of corn yield is the dependent variable. The land units to which the fertilizer is applied correspond to subjects in a psychological experiment. If the land units are homogeneous and differ from the land units that might be found in Nebraska or in another section of Iowa, there is no way to generalize the results obtained from this experiment to what might be expected in these other areas. Soil differences may make up important independent variables in the sense that they are related to the results obtained from the applications of the different amounts of the fertilizer. A particular amount of fertilizer may work very well with certain kinds of land and have little or no effect when applied to other kinds of land.

The potential errors involved in generalizing beyond a sample must obviously be kept in mind when attempting to apply the results obtained in a socio-psychological experiment with subjects from one culture to subjects from another culture.

The physiological processes of the human organism appear to be much the same regardless of whether the organism being studied is found among the Central Eskimo, the African Bantu, or the Australian Bushmen. An experiment which shows that a certain drug depresses or facilitates some physiological process in American subjects may be expected to result in similar reaction for subjects from other cultures. Is it possible that psychological processes such as perception, reasoning, learning and remembering, and so on, are also comparable among members of different cultures? If that were true, then it would be possible to project the results obtained from psychological experiments dealing with these processes to the members of cultures different from the one in which the experiment was originally conducted. The evidence on this point is surely not complete or decisive at this stage of psychological development, for relatively little research has been done on the problem of comparing psychological processes between members of different cultures.

The evidence available at the present time, however, does not indicate any reason to believe that the learning process of children and adults in the Northern United States differs from the learning process of children and adults in other sections of the United States. And, in general, this is true of other psychological processes. Hence, we may tentatively suggest that results of experiments relating to psychological processes will be applicable among all members of a given culture. But only additional experimentation can give a decisive answer to the problem.

It is often possible to incorporate into the formal experimental design one or more potentially relevant organismic variables. Thus, if age is a variable that we have reason to believe may be related to the observations of the dependent variable, the experimenter may take steps to see that subjects from different age levels are used in the experiment. If there is reason to believe that sex classification may be related to the outcomes, both men and women may be used as subjects. If sex is found not to be an important independent variable, then it is possible to generalize the results of the experiment for both sexes. On the other hand, if sex is found to relate to the outcomes of the experiment, we shall have two answers to the question investigated: one answer for males and another for females.

The importance of including organismic variables in the experimental design is that if these variables are not found to be relevant, the results of the experiment can be generalized on a much broader basis. They will be more generally applicable than would otherwise be the case. If the organismic variables do turn out to be relevant, then the experimenter will be in a position to suggest different answers to the problem he set out to investigate, depending upon the nature of the organismic variable involved. Further consideration will be given to this problem in a later section on experimental design.

NUMBER OF SUBJECTS

Let us assume that an experiment has reached the planning stage where the experimenter must give consideration to the number of subjects to be investigated or tested. How many subjects shall he use? There is no simple answer to this question and only suggestive answers can be given in terms of the objective of the investigation. If the objective of a given experiment is to be able to state rather precisely the degree of relationship between two variables, then it can be said that, in general, the experiment will require a larger number of subjects than would be the case if the objective were merely to determine whether any relationship at all exists between the two variables.

In the first case we are interested in estimating a parameter, the exact degree of relationship, within rather narrow limits. In the second case, we are interested in determining whether the parameter is greater than zero. In general, the amount of variability to be expected in estimates of parameters is a function of the size of the sample, i.e., the number of subjects investigated.

If the relationship that exists between the two variables is high, as measured, let us say, by a correlation coefficient, then it will require fewer observations to determine whether the relationship is significantly greater than zero than would be the case if the relationship is low or small. If the objective is merely to determine whether the relationship is significantly greater than zero, the number of subjects required will depend upon the degree of relationship that does exist. On the other hand, if we want to estimate the degree of the relationship rather precisely, a larger number of subjects will be required, regardless of whether the exact relationship is high or low.

In prediction studies, where the relationship between one or more independent variables and a dependent variable is studied and the results are then to be used for predicting performance of subjects not yet tested, a fairly large number of subjects is required. The reason is that the relationship found in the original study is used in the form of an equation to make predictions. If the relationship is not estimated rather exactly, the resulting predictions will be in greater error than would be the case with a more exact estimate. By means of statistical methods it is possible to estimate the number of subjects required to estimate parameters with a given degree of accuracy.

Consider another experiment, in which the objective is to determine whether the average measure on the dependent variable for one group is significantly higher than the average measure for another group on the dependent variable. In general, the number of subjects required in an experiment of this sort will depend upon the variability in performance of the subjects within each group and also upon the magnitude of the difference found between the two means. If we assume that the difference between the means remains constant, then the greater the variability in performance of the subjects within each group, the larger the number of subjects required in order to determine whether the difference between the two means is significant. If we can assume that the variability of performance remains constant within each group, then the greater the difference between the observed mean performance the fewer the number of subjects that will be required to detect this difference.

There are methods available for estimating the number of subjects required in an experiment of this sort. They require that we have available some estimates of the magnitude of the differences to be observed in mean performance and also some estimates of the degree of variability in performance to be expected. It is possible, also, that we may not have any interest in a difference unless it is of some specified magnitude. Given some estimate of variability of performance, it is then possible to estimate the number of subjects required in order to draw some conclusion concerning the value specified (Harris, Horvitz, and Mood, 1948). For further discussion of sample size see Operating Characteristics, and the Sign Test in Chapter 8.

Analysis of Variance

Consider the situation in which N subjects are divided at random into r groups of n subjects each. Assume that these subjects are all tested under exactly the same conditions, i.e., we have no independent variable operating, but only a measure of performance X. It is a typical finding that the observations made upon the subjects will vary. This variation can be measured by expressing each observation as a deviation from the mean m of all N observations, squaring the resulting deviations, and summing for all subjects. We shall call this the total variation or total sum of squares. This sum of squares may be written

$$\Sigma x_t{}^2 = \sum_{1}^{n} (X - m)^2. \tag{1}$$

The total sum of squares can, in turn, be analyzed into two component parts: one part will be based upon the variation of the individual n observations within each set about the set mean $\overline{X}$. This component of the total sum of squares measures the variation within groups about the group means and we may call it the sum of squares within groups. The second component of the total sum of squares will be based upon the variation of the r group means $\overline{X}$ about the mean m of all N observations. This component measures the variation in the group means and we shall call it the sum of squares between groups. The sum of squares within groups and between groups may be written

$$\Sigma x_w{}^2 = \sum_{1}^{r}\sum_{1}^{n} (X - \overline{X})^2, \tag{2}$$

$$\Sigma x_b{}^2 = \sum_{1}^{r} n\,(\overline{X} - m)^2. \tag{3}$$

Then

$$\Sigma x_t{}^2 = \Sigma x_w{}^2 + \Sigma x_b{}^2. \tag{4}$$

Dividing the sum of squares between groups and the sum of squares within groups by the appropriate number of degrees of freedom, $r - 1$ and $nr - r$, respectively, we obtain the mean squares, $s_b{}^2$ and $s_w{}^2$, shown in (5) and (6).

$$s_b{}^2 = \frac{\Sigma x_b{}^2}{r - 1}, \tag{5}$$

$$s_w{}^2 = \frac{\Sigma x_w{}^2}{nr - r}. \tag{6}$$

The analysis of a total sum of squares into a number of component parts, each with its own appropriate degrees of freedom, along with the tests of significance that may be made, is known as the *analysis of variance*. In the simple case described above, it can be shown that s_b^2 is an estimate of the variance of the individual measures, σ^2, *plus* n times the variance of the means, $n\sigma_{\bar{x}}^2$, whereas s_w^2 is an estimate of the variance of the individual measures only, σ^2. These two estimates may be written

$$s_b^2 \rightarrow \sigma^2 + n\sigma_{\bar{x}}^2 \qquad (7)$$

$$s_w^2 \rightarrow \sigma^2. \qquad (8)$$

TEST OF SIGNIFICANCE

In a typical experimental design, the r groups of subjects are tested under different experimental conditions constituting the independent variable. The experimenter's interest is in determining whether the experimental conditions relate to differences in mean performance of the subjects, i.e., whether the means may be considered to differ significantly. The test of significance is accomplished by comparing s_b^2, a measure of variation based upon the sum of squares between groups, i.e., variation based upon the average performance of groups of subjects treated differently, with s_w^2, a measure of variation based upon the sum of squares within groups, i.e., variation based upon the performance of groups of subjects treated alike.

The logic of the test of significance is determined by the consideration that if subjects have been randomly assigned to each of the r groups, and if the subjects have all been tested under exactly the same conditions, then the variation in the r group means should be the result of the process of randomization or chance factors only. On the other hand, if the subjects have been randomly assigned to the r groups and the groups have been tested under different experimental conditions, and if these are related to mean performance, then the variation of the group means will be increased, i.e., have a component which reflects the experimental conditions as well as random or chance factors. This component is indicated by the second term on the right in (7). Under conditions of random sampling and with no systematic differences in experimental conditions operating to influence the group means, the second term on the right in (7) will tend toward zero. Thus, on the average, the ratio s_b^2/s_w^2 should approach unity.

Suppose, however, that one of the experimental conditions operates systematically in such a way as to increase average performance, and that another experimental condition operates in such a way as to decrease average performance. Then, obviously, the various group means will differ more among themselves than would be the case if the experimental conditions contributed nothing to the mean performance. At the same time, if the influence of an experimental condition is systematic, i.e., if, in general, it tends either to increase performance or to decrease performance but not both, then the experimental condition will not tend to increase the variation within the group of subjects tested under this condition. Only in the rather atypical case where the experimental condition tends to increase the performance of some subjects and to decrease the performance of other subjects would the experimental condition contribute to the variation within groups. (For further discussion see Transformations in Chapter 8.)

In this typical experimental design, with n subjects assigned at random to r groups, the variation within groups provides a basis for evaluating the variation between groups. The variation within groups, s_w^2, is often called experimental error (Edwards, 1950b; Cochran and Cox, 1950; Fisher, 1942; Snedecor, 1946). The ratio, s_b^2/s_w^2, is the familiar F ratio, tabled by Snedecor (1946). In general, the larger the value of F, the smaller the probability that it is the result of chance, under the hypothesis being tested.

FACTORS CONTRIBUTING TO VARIATION BETWEEN GROUPS

We shall now consider some additional problems in terms of the factors that may contribute to the variation in performance of the subjects, and we shall see whether the influence of these factors is primarily upon the variation within groups (experimental error) or the variations between groups (experimental conditions). It should be obvious that if there are any factors contributing to the variation between groups, in addition to the experimental conditions themselves, then no clear interpretation of the significance of the experimental conditions will be possible. Any such factors will have to be controlled or randomized throughout all groups. In addition, it should be obvious that any factor which contributes to the variation within groups will make it more difficult to detect a significant variation in the group means. The measure of variation within groups, it will be recalled, provides the basis for evaluating the significance of the variation between groups. On the other hand, if experimental error, variation within groups, can be decreased, this, in turn, should add to the precision of the experimental design.

INDIVIDUAL DIFFERENCES

A major contributing source to the total variation in performance of subjects in an experiment is that of individual differences. It is found, for example, that with almost any test or measure made upon human subjects, the resulting observations will vary. Psychologists speak of this variation in terms of individual differences, regardless of whether the differences are acquired or innate. We have referred to such differences, characteristic of individuals, as organismic variables.

Consider an experiment in which the experimenter, *without randomization,* places subjects into one of r groups. For simplicity, let r be equal to two, and assume that one group is tested under one experimental condition A and the other is tested under a different experimental condition B. The experimental conditions constitute the independent variable, and we shall assume that the dependent variable is some measure of learning. The experimenter is interested in determining whether learning, the dependent variable, is in any way related to the independent variable, the experimental condition.

Suppose it is found that the A condition is superior to the B condition, as determined by the fact that the variation between the means of the two groups is significant when tested against the variation within groups, with the A mean being higher than the B mean. The experimenter draws the conclusion that the difference observed is a result of the independent variable, that is, the experimental conditions. However, let us assume that it is also found that the members of the A group have a significantly higher mean level of intellectual ability, as measured by some test of academic aptitude, than do the members of the B group. Since intellectual ability is probably positively related to learning ability, the experimenter has no way of knowing whether the superior performance of the A group under the experimental condition is the result of the experimental condition itself or is merely an artifact of the initial differences in the groups in terms of intellectual ability.

If the members of the A group are, in general, superior in intellectual ability to the members of the B group, this will operate as a systematic source of error in the measures of performance obtained under the experimental conditions. The mean of the A group, in other words, may be expected to be higher than the mean of the B group, regardless of the experimental conditions. Thus, if the means do differ on the dependent variable, this may be due either to an organismic variable, intellectual ability, or to the experimental variable, or both. The effect of the experimental condition is said to be *confounded* with intellectual ability and no clear conclusion concerning the experimental variable is possible.

RANDOMIZATION OF SUBJECTS

When organismic variables are not incorporated into the experimental design as independent variables whose influence is to be investigated, it is important to randomize these variables so that they will not have any systematic influence on the mean outcomes on the dependent variable. If subjects are randomly assigned to the different experimental conditions, then we may expect differences in organismic variables to be randomly distributed among the various experimental groups. The means of the experimental groups may thus be expected to differ as a result of the organismic variables only within the limits of random sampling.

Methods of randomly assigning subjects to experimental groups are described by Lindquist (1940), Walker (1943), and Edwards (1950b). With random assignment of subjects to the experimental conditions, organismic variables may be assumed to contribute to the random variation in the means of the groups, but not to any systematic variation in the means.

The necessity of randomization in the assignment of subjects may be made obvious if we consider an experiment in which 40 male and 60 female subjects are available. If an equal number of subjects is to be assigned to each of two groups, the method by which the subjects are assigned to the groups might result in having a preponderance of females in one group, with only a relatively small number of males. If the two groups of subjects are tested under different experimental conditions, and it is found that one group has a significantly higher mean on the dependent variable, this difference between the means may be the result of the experimental conditions or it may be the result of a systematic sex difference in performance, or both of these variables may contribute to the difference between the means.

In the case described, it would clearly be possible to assign the subjects so that the same number of males and females are present in each group. Balancing of subjects in this way may often be possible when the organismic variable is as obvious as it is in this case. But there will still be any number of other organismic variables operating, and on many of these the experimenter may have no knowledge of possible systematic differences which may be operative in performance under the experimental conditions. In what way shall the 60 females be divided into two groups of 30 each for the two experimental conditions? In what way shall the 40 males be divided into two groups of 20 each for the two experimental conditions? By randomly assigning subjects in each sex group to the experimental conditions, the experimenter has some assurance that these other organismic variables will be randomized and thus removed as a possible systematic source of error contributing to the variation in the means of the groups on the dependent variable. Failure to randomize the assignment of subjects means that the outcome of the experiment is always potentially open to more than one interpretation.

ARRANGEMENT OF CONDITIONS

Another systematic source of variation may occur with respect to the arrangement of the experimental conditions themselves. If, for example, one group of subjects is tested under one experimental condition at a particular time, in the morning, let us say, and a second group of subjects is tested under another experimental condition at another time, in the afternoon, let us say, then time of testing may contribute to any observed difference between the means of the groups on the dependent variable. Any systematic influence resulting from the difference in time of testing could, of course, be controlled by testing all subjects at the same time. Or the influence of time of testing could be randomized by the assignment of the experimental conditions in such a way that some subjects would be tested under each condition, in the morning and in the afternoon.

Time of testing itself might be introduced into the design as another independent variable to be investigated. This could be done by planning the experiment so that half of the subjects in each experimental condition are tested in the morning and half in the afternoon. Randomization in the assignment of the subjects would still be involved, however, in this design.

If prior evidence was available to indicate that time of testing did not contribute in any significant way to performance under the experimental conditions, this variable, of course, would not have to be considered in the plan of the experiment.

Temperature differences, differences in illumination, or differences in other environmental variables may be considered in the same way as the illustrative example of time of testing. All of these may be potential sources of systematic variation, to be randomized, controlled, or investigated.

IMPORTANCE OF RANDOMIZATION

The importance of randomization is stressed by Fisher in his discussion of the classical case of the lady tea-taster. This lady, he states, "declares that by tasting a cup of tea made with milk she can discriminate whether the milk or the tea infusion was first added to the cup" (Fisher, 1942, p. 11). The experimental design selected to investigate this claim is a relatively simple one: Eight cups of tea are prepared, four mixed in one way and four in the other. These cups are presented to the lady in random order for judgment. She has been told that there are four cups of one kind and four of the other. She is asked to taste the cups and to divide them into two sets of four each, in terms of how she believes the mixture to have been prepared. She is also told that the cups will be presented to her in random order.

The outcome of the experiment is evaluated in terms of the assumption that the lady does not have the ability to make the discrimination that she claims. If that is the case, the results of the experiment may be assumed to be governed by chance. Any observed outcome can then be evaluated in terms of the frequency with which that outcome would be expected in terms of the chance distribution. It can be shown that if nothing but chance is operating to influence judgments, then a completely correct classification of the cups would occur with a theoretical relative frequency of 1 time in 70.

If the lady makes a completely correct classification, this outcome would have a probability value of .014, under the hypothesis specified, i.e., that the judgments are a matter of chance. Consequently, the experimenter would reject the chance hypothesis for this outcome in terms of the usually accepted significance standard. Having rejected the chance hypothesis, can the experimenter conclude that the lady is correct in her claim? This is possible only if he has assurance that, lacking the discrimination, the completely correct classification actually does have a probability of .014. Fisher's comments on the necessity of randomization to ensure that this is the case are worth quoting in some detail:

> If all those cups made with the milk first had sugar added, while those made with the tea first had none, a very obvious difference in flavour would have been introduced which might well ensure that all those made with sugar should be classed alike. These groups might either be classified all right or all wrong, but in such a case the frequency of the critical event in which all cups are classified correctly would not be 1 in 70, but 35 in 70 trials, and the test of significance would be wholly vitiated. Errors equivalent in principle to this are very frequently incorporated in otherwise well-designed experiments.
>
> It is no sufficient remedy to insist that "all the cups must be exactly alike" in every respect except that to be tested. For this is a totally impossible requirement in our example, and equally in all other forms of experimentation. In practice it is probable that the cups will differ perceptibly in the thickness or smoothness of their material, that the quantities of milk added to the different cups will not be exactly equal, that the strength of the infusion of tea may change between pouring the first and the last cup, and that the temperature also at which the tea is tasted will change during the course of the experiment. These are only examples of the differences probably present; it would be impossible to present an exhaustive list of such possible differences appropriate to any one kind of experiment, because the uncontrolled causes which may influence the result are strictly innumerable. (Fisher, 1942, pp. 17-18)
>
> The element in the experimental procedure which contains the essential safeguard is that the two modifications of the test beverage are to be prepared "in random order." This, in fact, is the only point in the experimental procedure in which the laws of chance, which are to be in exclusive control of our frequency distribution, have been explicitly introduced. The phrase "random order" itself, however, must be regarded as an incomplete instruction, standing as a kind of shorthand symbol for the full procedure of randomisation, by which the validity of the test of significance may be guaranteed against corruption by the causes of disturbance which have not been eliminated. To demonstrate that, with satisfactory randomisation, its validity is, indeed, wholly unimpaired, let us imagine all causes of disturbance — the strength of the infusion, the quantity of milk, the temperature at which it is tasted, etc. — to be predetermined for each cup; then since these, on the null hypothesis, are the only causes influencing classification, we may say that the probabilities of each of the 70 possible choices or classifications which the subject can make are also predetermined. If, now, after the disturbing causes are fixed, we assign, strictly at random, 4 out of the 8 cups to each of our experimental treatments, then every set of 4, whatever its probability of being so classified, will certainly have a probability of exactly 1 in 70 of *being* the 4, for example, to which the milk is added first. However important the cause of disturbance may be, even if they were to make it certain that one particular set of 4 should receive this classification, the probability that the 4 so classified and the 4 which ought to have been so classified should be the same, must be rigorously in accordance with our test of significance.
>
> It is apparent, therefore, that the random choice of the objects to be treated in different ways would be a complete guarantee of the validity of the test of significance, if these treatments were the last in time of the stages in the physical history of the objects which might affect their experimental reaction. The circumstance that the experimental treatments cannot always be applied last, and may come relatively early in their history, causes no practical inconvenience; for subsequent causes of differentiation, if under the experimenter's control, as, for example, the choice of different pipettes to be used with different flasks, can either be predetermined before the treatments have been randomised, or, if this has not been done, can be randomised on their own account; and other causes of differentiation will be either (a) consequences of differences already randomised, or (b) natural consequences of the difference in treatment to be tested, of which on the null hypothesis there will be none, by definition, or (c) effects supervening by chance independently from the treatments applied. Apart, therefore, from the avoidable error of the experimenter himself introducing with his test treatments, or subsequently, other differences in treatment, the effects of which the experiment is not intended to study, it may be said that the simple precaution of randomisation will suffice to guarantee the validity of the test of significance, by which the result of the experiment is to be judged. (Fisher, 1942, pp. 19-21)

SYSTEMATIC ERRORS OF RATINGS

In some experiments different observers may be used to record or rate the behavior of subjects. Assume that two observers are available and that one observer has been assigned to each of two experimental groups whose performance is to be rated. If one of the observers tends to overrate systematically the performance of subjects in the group which he is observing, the mean performance score of this group will be higher than it would be otherwise, merely because of the rating tendency of the observer.

A comparison of the two means might reveal that one is significantly greater than the other. The experimenter may then conclude that the difference is the result of the independent variable, the experimental conditions, whereas the difference may be primarily a function of the observer and not the independent variable at all. Systematic errors of observation or rating of this kind will contribute to the variation in the means of the groups but not to the variability.

A suggested solution to the problem described might be to have both raters observe both groups and to pool their ratings, under the assumption that any systematic errors of rating would tend to operate in the same way for each of the two groups. If this assumption is valid, then the variance between groups would not be contributed to significantly by the errors of rating.

It is worth noting that systematic errors of rating will not be detected simply by correlating ratings of the same subjects obtained from two raters. Systematic errors of the sort described will have no influence upon the correlation coefficient. Two observers may show equal variability in their ratings of the same subjects and may also show perfect agreement in their ratings as measured by the correlation coefficient. Yet one observer may systematically over- or underestimate the performance of the individual members of the groups.

REDUCING EXPERIMENTAL ERROR

So far we have been discussing the necessity of controlling by experimental techniques or of randomizing systematic sources of error which might contribute to the variation between groups in addition to whatever variation might be contributed by the differences in experimental conditions. We shall now look at some factors contributing primarily to the variation within groups, i.e., to the measure of experimental error, and discuss the possible methods and techniques for reducing this source of variation.

RANDOM ERRORS OF MEASUREMENT

Let a group of N subjects be divided at random into r groups of n subjects each. Let the subjects in each group be tested under different experimental conditions and assume that the experimental conditions operate systematically to contribute to the variation in the means of the r groups on the dependent variable. We shall assume that no systematic sources of error are present, but that each single observation on the dependent variable is subject to random errors of measurement. By a random error of measurement we shall mean an error which shows no systematic tendency to result in over- or underestimates of the single observations. Stated in another way, we may say that an error of measurement is random if the probability that a single observation is overestimated by a given amount is equal to the probability that the observation is underestimated by the same amount.

In general, random errors of measurement will contribute to the variation within groups in such a way that the value of s_w^2 obtained will be larger than it would be if no random errors of measurement were present. As shown in (8), s_w^2 is an estimate of σ^2 and the test of significance of the variation in the experimental means will be made by dividing the estimate of (7) by that of (8) to obtain the F ratio. Thus

$$F = \frac{s_b^2}{s_w^2} \rightarrow \frac{\sigma^2 + n\sigma_{\bar{x}}^2}{\sigma^2}. \qquad (9)$$

An examination of (9) will show that if σ^2 is very large compared with $n\sigma_{\bar{x}}^2$ the value of F will not be greatly in excess of 1.0. If, for example, $n\sigma_{\bar{x}}^2$ remains constant and σ^2 is increased, the value of F approaches 1.0 as a limit. On the other hand, if the second term in the numerator of F does not vanish, then when σ^2 is decreased, the value of F will become very large.

Reduction of experimental error as measured by s_w^2 will thus enable the experimenter to detect more readily any systematic differences in the means of the experimental groups as measured by $n\sigma_{\bar{x}}^2$. Such reduction in experimental error can often be accomplished by increasing the reliability of the observations of the dependent variable and thus decreasing random errors of measurement.

INCREASING THE NUMBER OF REPLICATIONS

Let the difference between two means $\bar{X}_a - \bar{X}_b$ be D, let s_w^2 represent the variability of the individual observations about the group means and let n be the number of observations in each group. Then a test of significance of the difference, $\bar{X}_a - \bar{X}_b$, can be made in terms of the t distribution, where

$$t = \frac{\bar{X}_a - \bar{X}_b}{\sqrt{\dfrac{2s_w^2}{n}}}, \qquad (10)$$

s_w^2 is as defined by (6), and the usual assumptions of normality of distribution and homogeneity of variance are involved.

The value of t obtained by (10) may be compared with the table of t (Fisher and Yates, 1949) to determine its probability of occurrence under the hypothesis tested. In general, the larger the value of t, the smaller the probability that it is the result of chance or random variation.

It is apparent that if D is small relative to s_w^2, it will require a larger n to obtain a specified value of t than would be the case if D were large relative to s_w^2. In other words, the smaller the difference between two means with a constant value of s_w^2, the larger the number of observations that will be required to regard the difference as significant.

This same notion can be generalized to the case of more than two means. For this case, the test of significance may be made by (9). It is apparent in (9) also that if the true variation in a group of means, $\sigma_{\bar{x}}^2$, is small relative to the variation of the individual measures, σ^2, then a larger value of n will be required to obtain a specified value of F than would be the case if $\sigma_{\bar{x}}^2$ were large relative to σ^2. In other words, increasing n in (9) will result in a larger value of F, other things being equal.

In books on experimental design (Fisher, 1942; Cochran and Cox, 1950; Edwards, 1950b), increasing the number of observations in a given experimental plan is referred to as increasing the number of replications. Replication, in a sense, refers to the trials of an experiment. For example, if the experiment consisted of two experimental conditions with separate groups under each condition, one pair of subjects would provide one replication or trial of the experiment. Each additional pair of subjects would provide another replication of the experiment. With four experimental conditions and with a separate group of subjects tested under each condition, four subjects would provide one replication. Each additional set of four subjects would provide one additional replication.

From the discussion above it can readily be seen that increasing the number of replications of an experiment, for any specified true difference or variation in a group of means, will result in a larger value of t or of F than would be obtained with fewer replications. Since, in general, the larger the value of t or F observed, the smaller the probability that it results from random or chance factors under the hypothesis tested, it can be said that increasing the number of replications will increase the likelihood of detecting a given difference in the means of the experimental groups at a defined significance value.

SELECTION OF SUBJECTS

Since perhaps the largest contribution to s_w^2 in experimental work in psychology is the individual variation of the subjects, a reduction in s_w^2 can often be accomplished by the selection of subjects who are more homogeneous with respect to individual differences. Such selection may be based upon variables which the experimenter has reason to believe are correlated with the dependent variable.

Assume, for example, that an organismic variable is positively related to the dependent variable, as shown in Fig. 1. An examination of Fig. 1 shows that the subjects of high, aver-

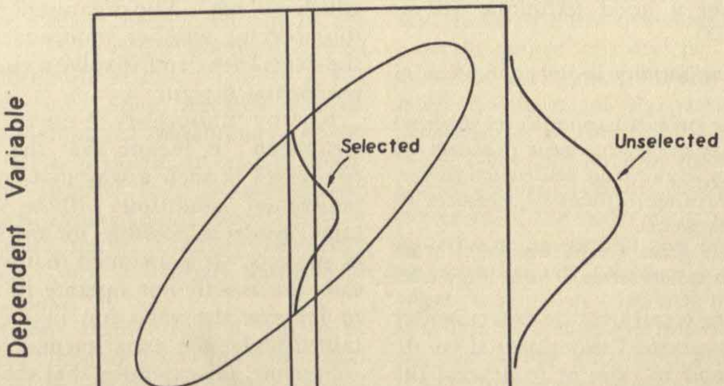


FIG. 1. The theoretical variation on a dependent variable of unselected and selected subjects on an organismic variable that is positively correlated with the dependent variable.

age, or low values of the organismic variable will vary less among themselves on the dependent variable than would be the case for unselected subjects covering the entire range of the organismic variable.

The experimenter may select subjects whose measures on the organismic variable place them in the average section. These subjects, in turn, may be assigned at random to the experimental conditions. If the experimental conditions contribute to the variation in the means of the groups, then this contribution will be more easily detected than would be the case with unselected subjects. This is because of the reduction in s_w^2 obtained by the selection of subjects on the organismic variable (it is assumed—and this will usually be the case—that the experimental conditions do not contribute significantly to the variation within groups).

The reduction in error, as measured by s_w^2, is achieved, in this instance, at some loss in the generality of the results of the experiment, assuming that the observed means of the experimental groups differ significantly. It would not be possible, for example, to generalize the results to subjects with either high or low levels of the organismic variables. They are not represented in the experimental groups and the results obtained may or may not hold for these subjects. In other words, without further research, the conclusions apply only to the level of the organismic variable selected and investigated in the experiment.

Experimental designs in which levels of an organismic variable can be used in such a way as to reduce the error variance without loss in generality of the conclusions drawn will be discussed in a later section.

COVARIANCE TECHNIQUES

Covariance techniques may also be used to reduce experimental error as measured by s_w^2. Covariance designs require that the experimenter have available measures on one or more organismic variables. If subjects are assigned at random to experimental groups and if the means are approximately equal with respect to the organismic variable for which measures are available, as they might be expected to be with random assignment, and if, at the same time, there exists a relationship between the measure of the organismic variable and the dependent variable, then the error variance will be reduced by taking this relationship into account. The amount of reduction in the error variance will depend upon the strength of the relationship that exists between the organismic variable and the dependent variable. Covariance analysis is described by McNemar (1949), Lindquist (1940), Cochran and Cox (1950), Johnson (1949), Edwards (1950b), and Fisher (1942).

EXPERIMENTAL TECHNIQUE

Experimental technique refers to the manner and conditions under which the experiment itself is carried out or conducted. Cochran and Cox (1950, p. 30) have pointed out that the main objectives of a good technique are as follows:

> i. To secure uniformity in the application of the treatments.
>
> ii. To exercise sufficient control over external influences so that every treatment produces its effects under comparable and desired conditions.
>
> iii. To devise suitable unbiased measures of the effects of treatments.
>
> iv. To prevent gross errors, from which no type of experimentation seems to be entirely free.

It is important to recall that the experimenter anticipates that a specified experimental condition will either tend to raise or to depress the measures obtained in a systematic fashion, so that the mean will be either raised or lowered, with no significant increase in variation of the individual measures. If an experimental condition tends to increase the performance of some subjects and to depress the performance of others, it is of importance to determine the associated organismic variable or variables which would enable the experimenter to predict which subjects will show improvement and which will not. The organismic variable would thus become another independent variable in the experiment and involve a change in the experimental design.

Now by "uniformity in the application of the treatments" is meant that the experiment is conducted in such a way that any specified experimental condition will be essentially the same, insofar as possible, for any specified group of subjects. It is assumed that when this is the case, errors will not operate in such a way as to increase the variation in the measures obtained under the experimental condition.

Suppose, for example, that subjects are tested individually under the experimental conditions and that the instructions to the subject are given by the experimenter. A difference in emphasis, stress upon one aspect, or other differences which might occur from subject to subject should be avoided, otherwise they may tend to provide "sets" for some subjects which will facilitate performance and "sets" for other subjects which will depress performance.

In psychological research, uniformity of the experimental conditions is of particular importance because of the subtleties of the interactions between experimenter and subjects. Subjects may react to the mood, appearance, and behavior of the experimenter as well as to the experimental conditions in which the experimenter is primarily interested. A friendly approach upon the part of the experimenter with some subjects and a more hostile or aggressive approach to others may introduce departures from uniformity which will influence individual performance.

Controlling "external influences so that every treatment produces its effects under comparable and desired conditions" is another aspect of a good technique. It is obvious that if external factors are not controlled with respect to the various experimental conditions, these factors may tend to have systematic influences upon the means. An experiment may be designed to test performance under conditions *A*, *B*, and *C*. The setup of the experiment may require that each test condition be assigned to a different experimental room. One of the rooms may differ from the other two in that it is poorly lighted and ventilated. Subjects tested in this room may show performance that is inferior to the performance of subjects tested in the other rooms. The difference observed in the means, however, may be the result of differences in the rooms rather than the result of differences in the experimental conditions which the research worker believes he is investigating.

Methods of recording and quantifying observations have already been discussed in another connection. The experimenter will need to know something about the reliability of the measures he obtains. Some research may require that new methods of measurement be developed—a research problem in and of itself. In addition, the changes in performance produced by different experimental conditions may be of such a subtle nature that the measuring instrument used is not sufficiently sensitive to reflect these changes.

Gross errors may result from a variety of causes. The experimenter may misread a dial or may miscalculate a score on a test. Careful observation will prevent many such gross errors, and examination of the records obtained from individual subjects may assist in the detection of gross errors. Measures obtained from single subjects which are out of line with those obtained from other comparable subjects may be the result of errors of observation, errors of recording, or errors of calculation.

> Faulty technique may increase the real experimental errors in two ways. It may introduce additional fluctuations of a more or less random nature. Such fluctuations, if they are substantial, should reveal themselves in the estimate of error as calculated in the analysis of variance . . . Where his estimated standard errors are consistently higher than those of other workers with similar material, the experimenter is advised to seek the reason, which may lie in differences in technique. In addition, faulty technique may result in measurements that are consistently biased. The estimate of error does not take account of such biases, since it is derived from comparisons of the measurements with one another; in other words, it estimates precision rather than accuracy. The principal safeguards against such biases are care and skill in the construction and handling of measuring devices, plus the intelligent use of randomization. (Cochran and Cox, 1950, pp. 30-31)

TESTS OF TECHNIQUE

Tests of technique can often be incorporated into an experimental design. This is accomplished by dividing the subjects at random so that each experimental group consists of a number of smaller random subgroups. Each set of subgroups will then constitute one trial of the complete experiment. A comparison of the results, that is, differences in the means, and of the error variances from the various trials of the experiment should provide some information concerning the technique of the experiment. If the separate error variances themselves show considerable variation beyond chance expectancy, the technique is open to question, since a good technique should not show significant variations in error upon repetition with comparable groups of subjects. A test of significance of variation in error variances can be made by means of Bartlett's (1937) test for homogeneity of variance.

Lack of consistency in the differences among the experimental means and in their ordering in the trials of the experiment would provide evidence with respect to the repeatability of the results obtained. Obviously, there is something wrong with an experimental technique in which comparable results are not obtained with comparable groups of subjects. On the other hand, if comparable results were obtained in the various trials of the experiment this would not necessarily reflect upon the adequacy of the technique and the absence of systematic sources of error. The same uncontrolled factors may be operating in each of the various trials, thus producing a spurious kind of consistency.

Functions of a Control Group

Many experiments in social psychology have been concerned with the influence of motion pictures, pamphlets, lectures, and other materials upon social attitudes. A typical experiment consists of giving a group of subjects an attitude test, usually called a pretest, then presenting the materials designed to influence the attitude, and finally giving the subjects a second attitude test, usually called a posttest. Solomon (1949) has called attention to the fact that many of the attitude studies summarized by Murphy, Murphy, and Newcomb (1937, pp. 946-979) have used this procedure.

The difference between the mean score on the pretest and on the posttest is found and tested for significance. If the means differ significantly, the experimenter concludes that the change in attitude is the result of the materials presented. This may or may not be the case.

It might be argued, for example, that the mere act of taking the first attitude test is sufficient to start subjects re-examining their attiudes and that this is the major factor producing the change. Any possible influence that the materials might have, in other words, is completely confounded with the potential influence of taking the pretest. The experimental design, in this instance, does not permit a clear interpretation of the results.

A more adequate experimental design involves a control group. The term *control group* in psychological research refers to a group of subjects whose performance on a dependent variable is to be used as a basis for evaluating the performance of one or more experimental groups on the same dependent variable. For the problem described, a group of N subjects could be divided at random into two groups of r subjects each. One of these groups would be designated the control group and the other the experimental group. Both groups would be given the pretest of attitudes. The experimental group would then be presented with the materials designed to influence their attitudes, but the control group would not. Both groups would then be retested. The experimental design may be summarized, following Stouffer (1950), as shown below:

Sequence	Experimental Group	Control Group
1. Pretest of attitude	$\bar{X}_1$	$\bar{X}_1'$
2. Materials presented	yes	no
3. Posttest of attitude	$\bar{X}_2$	$\bar{X}_2'$
Difference	$\bar{X}_1 - \bar{X}_2 = D$	$\bar{X}_1' - \bar{X}_2' = D'$

The difference D between the mean of the experimental group on the pretest $\bar{X}_1$ and on the posttest $\bar{X}_2$ is regarded as the change resulting from the taking of the pretest *and* the materials presented. If the amount of change shown by the control group is now subtracted from the change D shown by the experimental group, we have available a measure of the influence of the materials alone.

EXTENSION OF THE CONTROL GROUP DESIGN

In his discussion of the use of control groups in psychological research, Solomon (1949) has expressed his conviction that the design described above is still inadequate for a clear interpretation of the results. It is his contention that the pretest may *interact* with the experimental condition. Taking the pretest, he points out, "may conceivably change the 'set' or attentional factors important to the effectiveness of training" (p. 141) or materials presented. Or the pretest may influence the manner in which the materials are perceived by the experimental group.

To estimate the influence of possible interaction effects, Solomon proposes an extension of the control group design, consisting of the addition of a second control group. This second control group would differ from the first in that it would not be given the pretest, but would be presented with the experimental materials. The experimental design would be as shown below:

	Experimental Group	Control Group 1	Control Group 2
1. Pretest of attitude	$\bar{X}_1$	$\bar{X}_1'$	?
2. Materials presented	Yes	No	Yes
3. Posttest of attitude	$\bar{X}_2$	$\bar{X}_2'$	$\bar{X}_2''$
Difference	D	D'	D''

Since the second control group has not been given a pretest, the mean score for this group on the pretest must be estimated from the mean scores, $\bar{X}_1$ and $\bar{X}_1'$, obtained from the experimental group and the first control group. The difference D″ between this estimated mean and the mean on the posttest for the second control group is regarded as a measure of the influence of the experimental condition alone. The difference D' between the means of the first control group on the pre- and posttests is a measure of the influence of taking the pretest alone. The difference D between the means of the experimental group is a measure of the influence of taking the pretest, the experimental condition, *and* any interaction effects that may exist between the pretest and experimental conditions. An estimate of the interaction, designated by Solomon as I, is given by

$$I = D - (D' + D'') .$$

> There are many ways to interpret an obtained interaction. For example, if an interaction term is relatively large compared with the improvement of the experimental group, and if it is positive in sign, we might assume that the pretest somehow facilitated the effective use of training procedures. We then might start on a search for the kinds of attitudinal and perceptual changes that the pretest had produced, since such changes might have a generality of application for many kinds of educational procedures. In attitude testing, a negative I term would lead one to suspect that the pretest developed some kind of resistance to the propaganda material. This problem then could be approached by new experiments. It is possible that rules for building pretests which make subjects receptive to education might come out of such a program. In other words, the magnitude and direction of I ought to provide some hints to the experimenter. (Solomon, 1949, p. 143)

Additional extensions of the control group design, involving three control groups, are also discussed by Solomon (1949).

The use of control groups in psychological research will not always involve both a pretest and a posttest. A group of N subjects, for example, may be divided at random into two groups of r subjects each. One of these groups is designated as the experimental group and the other as the control group. Members of the experimental group experience the experimental condition (independent variable) and members of the control group do not. Both groups are then tested on the dependent variable. This design assumes that nothing occurs in the time interval between the division of the subjects at random into the two groups and the time of testing on the dependent variable that would differentially influence performance of the two groups on the dependent variable. Nothing, that is, except in the case of the experimental group, which is, in the interval, put through the experimental condition.

Designs involving both a pre- and posttest of the experimental and control groups have been called by Stouffer (1950) and Hovland, Lumsdaine, and Sheffield (1949) *before-after* designs. When the experimental procedure involves only the posttest, this has been called an *after-only* design. The relative merits of these two kinds of designs are discussed in some detail from the point of view of statistical and sampling consideration, the kinds of analyses of the data possible, and the possible effects of the measuring process on the measurements observed, by Hovland, Lumsdaine, and Sheffield (1949, pp. 308-328). In addition, these authors provide many examples of applications of both types of design in research on problems of mass communication.

NUMBER OF SUBJECTS TO BE ASSIGNED TO CONTROL AND EXPERIMENTAL GROUPS

If subjects have been divided at random and we may assume that the variation within the two groups will be approximately the same on the dependent variable, then it can be shown that the most efficient division of the subjects is such that we have an equal number in each group. In general, "we may denote as an *efficient design that distribution of the total population among the treatment groups which will yield maximally significant differences when the comparisons are made between each experimental group and the appropriate control group*" (Kurtz, Jaspen, and Ash, 1951, p. 580).

If two experimental groups are used along with one control group, the most efficient division of the subjects is such that each experimental group has 29.29 percent of the available subjects, with 41.42 percent in the control group. Tables have been prepared by Kurtz, Jaspen, and Ash (1951) showing the most efficient division of the subjects for one to four control groups and from one to ten experimental groups. The tables are so constructed as to show the division of subjects when the assumption of homogeneity of variance is not tenable, as well as for the case where this assumption is tenable. These tables are designed for use when the comparisons to be made are between the mean of the control group and each of the respective means of the experimental groups.

Randomized Groups Design

One of the most frequently used and simplest experimental designs is called the *randomized groups* design. In this design N subjects are divided at random into r groups of n subjects each. It is not necessary that the groups have equal n's, although the analysis of the data is somewhat simpler when this is the case. One of the groups may be a control group and the other groups may be experimental groups. For each subject we have available a measure *X* on the dependent variable. The independent variable consists of the different experimental conditions.

The randomized groups design is of value in problems where the differences among a set of experimental means are to be evaluated or where comparisons are to be made between a control group and each of the experimental groups.

A randomized groups design, for example, might be used to investigate:

> . . . the influence of different periods of food deprivation upon learning; the influence of different numbers of reinforcements upon conditioning and extinction; the influence of different methods of instruction upon achievement; the influence of different sets of verbal instructions upon problem solving; the influence of different sensory cues upon maze learning; the influence of different kinds of motivation upon performance; the influence of different periods of rest upon fatigue; the influence of different kinds of interpolated activities upon learning; the influence of different kinds of work situations upon production and fatigue; the influence of different periods of practice upon learning. (Edwards, 1950b, p. 185)

The randomized groups design has the property of having the error variance based upon a larger number of degrees of freedom than any other design using the same number of subjects. This is likely to be of value in experiments with a small number of observations, where a possible reduction in the error variance by an alternative design would not be offset by a corresponding loss in the degrees of freedom for the error variance. At the same time, the error variance will include all of the variation attributable to individual differences. When subjects are fairly heterogeneous in performance on the dependent variable, the design may require an excessive number of subjects in order to detect significant differences in the means of the experimental groups.

RANDOMIZED BLOCKS DESIGN

In *randomized blocks design* subjects are divided into a number of blocks or groups, each of which, the experimenter has reason to believe, will be more homogeneous in performance on the dependent variable than unselected subjects.

The term "randomized blocks" comes from agricultural experiments in which the experimental unit is a plot of land, and a block corresponds to a section of land consisting of adjacent or neighboring plots. There is reason to believe that plots of land which are near to one another are more alike with respect to fertility and general soil condition than plots which are more remote. In psychological research, the experimental unit is the subject instead of the plot, the relatively homogeneous group corresponds to a block, and individual differences correspond to fertility and general soil condition. Each block will, in general, consist of *r* subjects, where *r* is the number of experimental conditions. With *n* such blocks available, the experiment will be replicated *n* times.

On what basis does the experimenter make up the blocks or groups of subjects? This will depend a great deal upon the insight of the experimenter and upon his knowledge of the dependent variable and related organismic variables. If he knows, for example, that a particular organismic variable correlates substantially with the dependent variable, the organismic variable can be made the basis of the subdivision of the subjects into blocks. An examination of Fig. 1 will show that if vertical slices are made through the correlation diagram, the subjects thus selected within each slice will be more alike with respect to the organismic variable and that they will also be more alike with respect to performance on the dependent variable.

Assume that *r* experimental conditions are to be tested and that the experimenter has measured *N* subjects, where *N* is some multiple of *r*, with respect to some organismic variable that he believes is positively related to performance on the dependent variable. If the subjects are arranged in order of magnitude of their values on the organismic variable, then the first *r* subjects will make up the first block or group, the next *r* subjects the second block or group, and so on, until *n* such blocks or groups have been selected. All that now remains is for the experimenter to assign the subjects within each block at random to the various experimental conditions, with one subject in a block to each condition.

Since, in psychological research, the subjects will be *matched* with respect to an organismic variable within each block, the randomized blocks design has also been called a matched groups design (Edwards, 1950b, pp. 264-283). Each group or block of subjects may be considered as representing a level of the organismic variable. If the matching has been successful, in the sense that the levels of organismic variable used for matching the subjects are positively and substantially correlated with performance of the subjects on the dependent variable under the experimental conditions, then, in general, the error variance for the matched groups design will be less than the corresponding error variance for the randomized groups design. The reason is that some of the variation included in the error variance of the randomized groups design will now be attributable to the variation in the means of the replications or blocks.

The randomized blocks design is also applicable in the case of two groups, where one group serves as a control group and the other as an experimental group.

In the matched groups design discussed, it has been assumed that each of the r groups will be representative of a single common population described by the matching variable. That is, the means and variances within the several groups should be approximately equal with respect to the matching variable. The procedure of taking a block of homogeneous subjects and then assigning one member from the block to each of the experimental conditions at random, pretty much ensures that the groups will be comparable within the limits of random sampling errors. The independent variable in this design will be the differences in the experimental conditions to which the subjects are assigned and the dependent variable will be the performance measures of the subjects obtained under the experimental conditions.

REGRESSION EFFECTS IN MATCHED GROUPS DESIGNS

There are other cases of experimental designs using matched groups (Thorndike, 1942) which involve additional considerations. These considerations concern the phenomenon of *regression*.

Assume that the distribution of measurements on an X variable is as shown in Fig. 2. We assume that this variable is positively related to another variable Y, but that the relationship is not perfect. If we select a group of subjects who are all one standard deviation above the mean on the X variable and another group of subjects who are all one standard deviation below the mean on this variable, the difference between the means of these two groups, in standard deviation units, will be 2.0. If it is known that the correlation between the X variable and the Y variable is .80, then it can be predicted that if the two groups are tested on the Y variable, without any intervening experimental conditions, the difference between the means on this variable will be less than 2.0 standard deviation units or, more precisely, that the difference will be 1.6 standard deviation units.

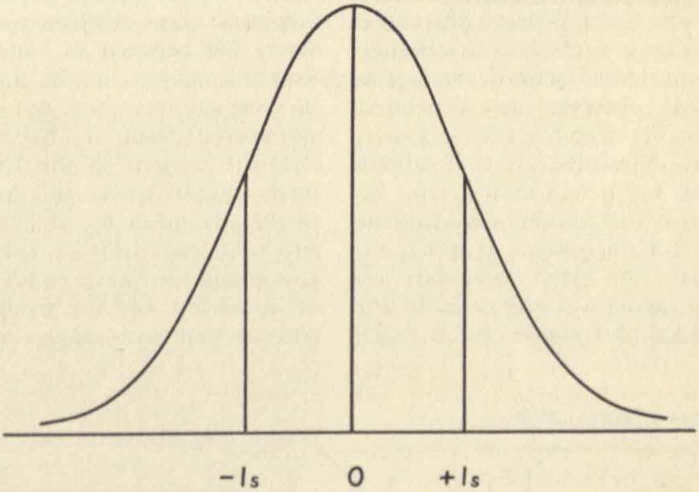


FIG. 2. Normal distribution with mean equal to zero and standard deviation equal to 1.0.

This prediction is based upon the phenomenon of regression. The scores of both groups of subjects will regress toward the common population mean, which, for scores in standard deviation units, is equal to 0. Thus the mean score of the group which was 1.0 standard deviation units *above* the mean on the X variable, may be predicted to be only .8 standard deviation units above the mean on the Y variable. The mean score of the group which was 1.0 standard deviation units *below* the mean on the X variable may be predicted to be only .8 standard deviation units below the mean on the Y variable. The scores for each group will regress, move back toward, the population mean.

The degree of regression that is present on the Y variable can be shown to be inversely related to the coefficient of correlation between the X and Y measures and directly related to the amount of deviation of the X measures from the X mean.

Now, assume that the X variable is a variable upon which subjects are to be matched and that the Y variable is a dependent variable. It should be clear that the procedures described for matching in the case of a randomized blocks design will, in general, tend to assign subjects to the experimental groups in such a way that the members of each group will show comparable deviations from the X mean. For each subject who is a given distance above or below the mean on the X variable in any one of the groups, there will be a corresponding subject in each of the other experimental groups. Consequently, in this situation, regression should influence each of the groups in much the same way. In the randomized blocks design the primary purpose of matching of subjects has to do with achieving a reduction in the error variance of the dependent variable.

We now consider a case, described by Thorndike (1942), in which the primary purpose of matching subjects on a variable is to eliminate this variable as a potential factor in the explanation of the results observed on a dependent variable. Assume, for example, that an experimenter wished to determine whether subjects taking a course in Latin will show greater improvement in English vocabulary than students who do not take such a course. It would not do just to compare the mean vocabulary test scores of a group taking a course in Latin and a group not taking this course, for it might well be argued that the two groups differed initially with respect to their vocabulary test scores as well as possible other variables related to academic performance.

Assume that it is true that students who propose or plan to take a course in Latin constitute one population with respect to vocabulary test scores and let the mean of this population be designated as m_2. Assume also that students who do not propose to take a course in Latin constitute another population with respect to vocabulary test scores and let the mean of this population be designated as m_1. We shall assume that the distributions of vocabulary test scores for a large number of students in each population are as shown in Fig. 3.

Obviously, if students taking a course in Latin are tested for vocabulary only at the conclusion of the course and their mean is compared with the mean vocabulary test score of students not taking a course in Latin, we may expect the two means to differ on this variable. The course in Latin may contribute little or nothing to the observed mean difference, which may merely be the result of initial differences in vocabulary of the two groups. To control initial differences in vocabulary, the experimenter decides to match two groups upon the basis of their initial vocabulary test scores.

Examination of Fig. 3 will indicate that if individuals in the non-Latin group are to be matched with individuals in the Latin group, most of the successful matchings will occur with respect to those subjects in each group whose scores fall between m_1 and m_2. In general, then, the subjects in the non-Latin group will be those subjects whose test scores are *above* the population mean m_1 for non-Latin students. And the subjects in the Latin group will be those subjects whose test scores are *below* the population mean m_2 of Latin students. With successful matching of pairs of subjects, the two groups will have *equal* means on the first or initial test, and the groups might be said to be equivalent with respect to this variable. We

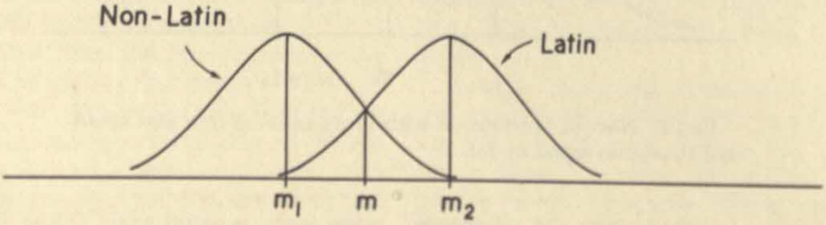


FIG. 3. Theoretical distributions of vocabulary test scores for students planning to take a course in Latin and for students not planning to take a course in Latin.

shall indicate this common initial mean as m.

But consider now the influence of regression in the absence of any experimental condition. If the two groups are retested on another form of the vocabulary test Y, which is to be used as a measure of the dependent variable, what may we predict? Even without the introduction of any experimental condition, we may predict that the mean score for the Latin group on the second vocabulary test will be higher than that of the non-Latin group. The reason for this is that since the correlation between X and Y may be high, but not perfect, the Latin students will show regression in their test scores on the second test toward the mean of the population from which they were drawn, i.e., m_2. In other words, on the second test the mean should move away from m and toward m_2 for the Latin students. And for the non-Latin group, we may predict that the mean m will be displaced toward m_1, the mean of the population from which they were drawn. It is important to note that the regression of scores is, as Thorndike (1942, p. 94) points out, *"toward the mean of the particular population to which they belong."*

The differential influence of regression, in this instance, might account for the difference observed between the means of the two groups of subjects on the second test, and the Latin course itself may contribute little or nothing to this observed difference. The problem is to obtain two groups who may be expected to be equivalent on the second test, apart from any difference attributable to the course in Latin, not two groups who are equivalent on the first test. Thorndike suggests the following solution:

> We should match members from our two groups in terms of *predicted final test score.* Again the regression equation for predicting final test score will be different for each population and each final test score must be predicted in terms of the regression equation for that population. The regression equation for each population must be the one which holds *when the special experimental treatment is not applied,* or else the effect of the experimental treatment will be absorbed into our regression equation. (Thorndike, 1942, p. 101)

Using the basis for matching suggested above, the observed difference between the means of the two groups on the second test may be assumed to be independent of any differential regression influence.

Comparable problems would arise if subjects high and subjects low on a measure of intelligence were matched with respect to another measure of ability, say verbal ability. If one of these two groups was then given some kind of experimental treatment and retested with a second test of the ability measure, we would undoubtedly find that the two means now differed on the matching variable. This observed difference also would be an instance of the regression phenomenon, and not necessarily the result of any differential influence of the experimental condition upon the verbal ability of bright and dull subjects.

Ex Post Facto Design

Another kind of matched groups design which has had frequent use in social psychology has been described in detail by Chapin (1947). Chapin calls the design an "ex post facto" study. The design may be illustrated by Christiansen's (1947) study of the economic adjustment of high school graduates and non-high school subjects.

We may designate the high school graduates as the experimental group and the non-high school graduates as the control group. The factor of interest was the economic adjustment (dependent variable) of these two groups as related to educational status (independent variable). Since differences between the two groups on variables other than education might be expected to relate to economic adjustment, an attempt was made to control these variables by matching the members of the experimental and control groups on as many of the variables as possible.

When the two groups were matched by equating means and variances on five factors, the initial 1130 members of the high school groups were reduced to 200. Matching on a sixth factor reduced the number of subjects from 200 to 145. When individuals were paired in the matching, rather than simply striving for equal means and variances on each of the matching factors, the six factors resulted in only 23 successful matches.

> The expectation, of course, in the ex post facto design is that factors which are equated for the two groups will not influence the comparison on the variable of investigative interest. That this may not be the case when a large number of subjects are eliminated from the experimental group because successful matches cannot be made, is demonstrated by a study of Freedman and Hawley (1949). Freedman (1950) reports, for example, that in comparing educational differences between matched groups of migrants and nonmigrants, a significant difference was observed

> when 58 percent of the original group of migrants was matched with a control group of nonmigrants, but the difference was not significant when 89 percent of the original group was matched. In other words, a difference which may be present between subjects easily or readily matched may not be the same as the difference observed between subjects who are more difficult to match. (Edwards, 1951a, p. 336)

In addition to the problem mentioned above, the ex post facto study design is subject to regression complexities described earlier. It seems likely, for example, that the high school subjects in Christiansen's study constitute one population with respect to economic adjustment and the non-high school subjects constitute another population.

In a study such as Christiansen's, the research interest is in the relationship of the group classification and a dependent variable, when subjects are truly equivalent on some matching variable or background trait. No experimental conditions are involved or implied in the comparison of the groups on the dependent variable. Thorndike (1942, pp. 100-101) points out that in this instance:

> If the groups are to be equivalent in true score on the background trait, they must be matched on the basis of *predicted true score* — i.e., score predicted by the regression equation between original test on the background trait and a retest at the time of the experimental comparison. Since the regression equations for the different populations *will not be the same* in the case which we are now considering, the predicted true scores for each individual must be determined from *the regression equation for his own population.* Groups matched in this way will be truly equivalent upon the background trait, and differences between them in the experimental trait must be due to some factor other than background trait differences.

FACTORIAL DESIGN

In some experiments, we may have two or more independent variables, each of which is varied in two or more ways, sometimes called *levels.* If we make all possible combinations of these variables and incorporate each combination as a separate experimental condition, we have what has come to be known as a *factorial design.*

Consider the case of two variables, designated by the letters A and B. Let A be varied in two ways which we may designate A_1 and A_2. For simplicity, we shall let B be varied in only two ways also, B_1 and B_2. In general, if we let the subscripts i and j stand for the number of ways in which each of the variables is varied, we shall have $i \times j = r$ experimental combinations, each combination making up one of the experimental conditions under which subjects are to be tested. For the example described, we have $2 \times 2 = 4$ different combinations. These combinations are shown below:

A_1B_1	A_2B_1
A_1B_2	A_2B_2

One replication of the experiment described will require r equals 4 subjects. If we have N subjects available, where N is some multiple of r, then the subjects may be divided at random into r groups of n subjects each. With 80 subjects, for example, we would have 4 groups of 20 subjects each. This factorial design would permit a test of significance for the *main effects,* i.e., the differences between the A means and the differences between the B means, and also for the *interaction* that may be present between A and B.

The concept of an interaction was introduced earlier in connection with control groups. To see more precisely what an interaction measures in the present case, assume that the A variable represents the number of presentations of a transcription of a short lecture, with A_1 corresponding to one presentation and A_2 to two presentations. Let the B variable represent whether the transcription is read by a male, B_1, or a female, B_2. We shall assume that 80 male subjects have been divided at random into 4 groups of 20 subjects each. One group is then assigned at random to each of the 4 experimental conditions. The dependent variable is a measure of learning determined by a test covering the material presented in the transcription. The mean scores on the dependent variable of the four experimental groups are as shown below:

Group:	A_1B_1	A_1B_2	A_2B_1	A_2B_2
Mean:	58.0	34.0	72.0	64.0

The following questions may be raised: Are two presentations superior to one? Is the male speaker superior to the female? Is there any interaction between the number of presentations and the sex of the speaker? The four means may be arranged in a 2×2 table as follows:

	B_1 Male	B_2 Female	Mean	$B_1 - B_2$
A_2 Two presentations	72.0	64.0	68.0	8.0
A_1 One presentation	58.0	34.0	46.0	24.0
Mean	65.0	49.0	57.0	
$A_2 - A_1$	14.0	30.0		

To determine whether the number of presentations is a significant variable, we would test for the difference between the means 68.0 and 46.0. To determine whether the sex of the speaker is a significant variable, we would test for the difference between the means 65.0 and 49.0. Now, if there is *no* interaction between the two variables, then we should expect the difference in the amount learned under a male and female speaker to be the same, within random errors, regardless of the number of presentations. In this instance, the difference between the means of the male and female speakers for one presentation, 24.0, and the difference between the means of the male and female speakers for the two presentations, 8.0, should estimate the same quantity, i.e., the difference due to the sex of the speaker. The average of these two means would be $(8.0 + 24.0)/2 = 16.0$ and represents the main effect of the *B* variable. It can be observed that this value is the same as the difference between the two means for the male and female speakers referred to earlier, i.e., 65.0 — 49.0 equals 16.0.

On the basis of the same argument, assuming no interaction, the difference between one and two presentations for the male speaker, 14.0, and the difference between one and two presentations for the female speaker, 30.0, should estimate the same quantity, i.e., the difference attributable to the number of presentations. The average of these two values would be $(14.0 + 30.0)/2 = 22.0$ and it can be observed that this is the difference between the means for one and two presentations referred to earlier, i.e., 68.0 — 46.0 equals 22.0.

It seems apparent in the present instance, however, that an interaction effect is present. The basis for this observation is the discrepancy between the values 8.0 and 24.0, which supposedly estimate the same quantity, i.e. the difference attributable to the sex of the speaker. It seems obvious that the superiority of the male speaker is much less when two presentations of the transcript are used than when only one is used. The degree of superiority of the male speaker over the female speaker, in other words, is influenced by the variable "number of presentations," and we call this an interaction effect. A measure of the interaction effect is given by the absolute difference between 8.0 and 24.0, which equals 16.0.

Now examine the two values 30.0 and 14.0, which we previously assumed to be estimates of the same quantity, the influence of the variable "number of presentations." It is apparent from the discrepancy between these two values that the superiority of two presentations over one presentation is much greater in the case of the female speaker than in the case of the male speaker. In other words, two presentations seem to have a much greater influence in improving learning over one presentation in the case of the female speaker, and to contribute much less in the case of the male speaker. Again, we call this an interaction effect between the two variables, and this interaction is measured by the absolute difference between 14.0 and 30.0, which is equal to 16.0. This is the same interaction that we discussed in the previous paragraph. We have simply gone at it from the point of view of the *A* variable rather than of the *B* variable.

It is important to note that in the present experiment it is impossible to generalize beyond the particular male and female speakers who made the transcript. It is not, in other words, possible to conclude that male speakers, in general, will be superior to female. No random sample of speakers of the two sexes has been incorporated into the design. Similarly, we cannot generalize beyond the one and two presentations tested in the experiment. And it is clear, if the interaction effect is significant, that any generalization concerning two presentations versus one will be qualified in terms of the interaction between number of presentations and sex of speaker. It is equally true that the generalization concerning sex of the speaker will be qualified in terms of the interaction of this variable with the number of presentations.

The advantage of the factorial design in research is that it enables the experimenter to investigate several variables in all possible combinations. If only the main effects or a portion of them turn out to be significant, the conclusions drawn will be established on a broader base than if the variables were studied one at a time in isolation from all other variables. At the same time, information is gained with respect to several variables rather than with only one variable and this is accomplished with an economy of observations.

If interaction effects are significant, information is available about the nature of the interactions of the variables, as is clear from the illustrative example. For these reasons, the factorial design is apt to be of great value in the exploratory stages of research, where little is known about the possible ways in which several independent variables may interact to produce changes in the dependent variable.

LATIN SQUARE DESIGN

If r experimental conditions are to be investigated, these can be arranged in the form of an $r \times r$ square in such a way that each experimental condition appears once and only once in each row and once and only once in each column. If we let the letters, A, B, C, D, and E correspond to the different experimental conditions, then an example of a 5×5 Latin square is shown below:

$$\begin{matrix} B & E & D & C & A \\ C & A & B & E & D \\ D & B & C & A & E \\ E & C & A & D & B \\ A & D & E & B & C \end{matrix}$$

The assignment of the letters to the cells is at random, subject only to the restriction that each letter appears once and only once in each row and each column.

In the Latin square design, as used in psychological research, the rows usually correspond to subjects. Each subject, for example, is tested under all of the experimental conditions but in a different sequence or order. The columns of the square show which particular experimental condition is to be applied to a given subject at a given time or period. It will be noted that every row (subject) represents a complete replication of the experiment and that every column (trial or period) also represents a complete replication of the experiment. The purpose of this double grouping is to eliminate from the total variance the variation attributable to differences among the row means and differences among the column means. The remaining variation can then be analyzed into two component parts: that attributable to differences among the means of the experimental conditions and a residual component called experimental error.

For psychological problems, the Latin square design should be most useful when the number of experimental conditions ranges between 4 and 9. With fewer than 4 experimental conditions, the reduction in the degrees of freedom available for the error variance may offset any gain resulting from the control of the row variation and column variation. With more than 9 experimental conditions the problems involved in repeated testings of the same subject may prove to be difficult.

More complex designs involving the Latin square can be arranged. It is possible, for example, to replicate the same Latin square with an additional set of r subjects. The analysis of the separate Latin squares will then provide a test of technique as described earlier, as well as a test of the particular orders or sequences of the experimental conditions involved. It is also possible to provide replications of the complete experiment through the use of additional sets of r subjects for which the arrangement of the experimental conditions is determined by other randomized Latin squares which differ from the original square. Procedures for pooling and analyzing the observations obtained from these designs are described by Edwards (1950a) and Grant (1948). Other variations in the Latin square design are discussed by Cochran and Cox (1950).

Some disagreement has centered about the use of the Latin square design in psychological research. The tests of significance involve the assumption that interaction effects are not present, and McNemar (1951) has stated his belief that this assumption is not tenable. The possible interactions are those involving subjects, trials or test periods, and experimental conditions. The Latin square design, however, has been extensively applied in biological research and in other fields and it is somewhat difficult to believe that the conditions of application in psychology are vastly different from those in biology.

From the experimenter's point of view, the major question would seem to be this: If one particular experimental design is used and certain conclusions are drawn concerning the differences observed among the experimental means, would these conclusions be changed upon repetition of the experiment using another experimental design? It is undoubtedly true that some differences will be observed with respect to the probabilities attached to the tests of significance, yet, at the same time, it seems unlikely that with a good experimental technique the conclusions drawn from two different designs would be reversed. This is a matter, of course, that could be put to empirical test.

Balanced Latin Square Design

In some experiments in which a Latin square design is to be used, it may be desirable to arrange the square so that the experimental conditions are balanced with respect to the immediately preceding and following conditions. Bugelski (1949) cites some cases where this may be so. In a balanced Latin square, each experimental condition is preceded and followed by every other experimental condition equally often. The following is an example of a 4 × 4 balanced Latin square.

A B C D
B D A C
C A D B
D C B A

Taking successive pairs of experimental conditions, we have: *AB, BD, CA, DC, BC, DA, AD, CB, CD, AC, DB,* and *BA.* Rearranging these pairs of experimental conditions gives

AB AC AD BC BD CD
BA CA DA CB DB DC

and it may be noted that each condition precedes and follows every other condition but once. This design would tend to minimize any systematic tendency for one condition to increase or depress performance upon a subsequent condition. It is possible to construct a balanced Latin square of the sort described for any situation where the number of experimental conditions is an even number.

When the number of experimental conditions is an odd number, a balanced design may be arranged with two Latin squares. For this balanced design, each condition will precede and follow every other condition twice. An example involving 5 experimental conditions is shown below:

A B E C D
B C A D E
C D B E A
D E C A B
E A D B C

D C E B A
E D A C B
A E B D C
B A C E D
C B D A E

Methods of constructing balanced Latin square designs for both an even and an odd number of experimental conditions are described by Edwards (1951b).

Conclusion

This chapter has touched briefly upon the problems of experimental design: the planning and execution of experiments. The principle of randomization, the nature of replication, methods of reducing experimental error, experimental technique, systematic errors, tests of significance, selection of subjects, and the general problem of the design and analysis of experiments are treated in greater detail in Fisher (1942), McNemar (1949), Snedecor (1946), Edwards (1950b), Tippett (1941), Goulden (1939), Cochran and Cox (1950), Dixon and Massey (1951), Mather (1947), Johnson (1949), and Kempthorne (1952).

Although general principles of experimental design can be discussed in books, books themselves cannot provide the answers to, or even make one aware of, the specific problems of research that one encounters anew each time a research problem is undertaken. This ability to deal with specific problems is something that perhaps can only be learned by doing, i.e., through an intensive period of on-the-job training, at first under the supervision of a skilled research worker, and then, later, through years of individual research.

The beginning research worker in psychology can bring to his first research project a thorough knowledge of statistical analysis and the principles of experimental design. He should also possess a good background in the techniques and methods previously developed for the making and recording of observations. He also needs a sound basis in the factual material of his subject matter specialty, and some degree of theoretical sophistication is also an asset. And finally, he should come with a thorough knowledge of the organisms he plans to investigate, whether they be mentally deficient or normal children, normal adults or abnormal patients, white rats, or cocker spaniels. But all of this knowledge cannot be a substitute for actual experience in research. It is only by way of preparation for research.

That on-the-job training for research is desirable is recognized by most major departments of psychology in universities which specialize in the training of psychologists. One of the primary purposes of the master's thesis is to provide a research problem that is ordinarily carried out under the supervision of a senior adviser. In those relatively few departments which do not require a master's thesis, on-the-job training in research may be provided for through minor research projects which are undertaken as part of the traditional academic course work.

After three to five years of graduate academic training, the future research psychologist usually undertakes his first major individual research project, his doctoral dissertation. His first concern at this time is the selection of the problem to be investigated. Psychology is a relatively young science, and the problems for which we require answers are still more numerous than the problems for which answers are already available. Yet, the selection of a research problem is a task that has caused many graduate students a prolonged period of anxiety and concern.

As has been pointed out elsewhere (Edwards, 1946) an obvious and excellent source of research problems can be found in the theories advanced by various writers in a given field. In the psychology of learning, for example, the theories of Hull (1943), Tolman (1932), Guthrie (1952), Skinner (1938), and Miller and Dollard (1941), have been the starting point of many experiments. Sears (1943) has summarized many experiments which had their origins in psychoanalytic theory. The exposition of Gestalt theory by Koffka (1935) and the variations of Gestalt theory proposed by Lewin (1935, 1936) have resulted in many important research projects. The theoretical writings of Allport (1937) and Murray (1938) in the field of personality are still other examples of a rich source of research problems.

> Students who have been indoctrinated with critical attitudes and who read challengingly the literature in their field of specialization will not fail to find innumerable projects demanding investigation. Students who read with blind acceptance of authority and the printed word, on the other hand, are apt to overlook assertions and statements made by writers which are not based upon research evidence. Such assertions and statements, when questioned, may very well become the basis of research problems for the critical reader. The proverb "The grass looks greener on the other side of the fence" and the many variations of this proverb, when questioned, became the basis of a series of ingenious experiments for Wright (1937). And Irwin's (1943) questioning attitude applied to certain aspects of Wright's research led to additional experimentation. (Edwards, 1946, pp. 278-279)

As has been emphasized in the introductory sections of this chapter and again in the above discussion, the posing of a question is the first step in research. A question, properly phrased, becomes a hypothesis which can then be subjected to empirical test. After a question has been formulated, the next step involves the planning of the research. This planning consists of determining what data will need to be collected, the manner in which they will be collected, and the methods by which they will be analyzed. Many hours go into the planning stage of an experiment, but these hours are well spent. It is at this time that consideration is given to the controls that will be necessary, the construction of the apparatus that may be needed for the experiment, the technique and method by which the observations are to be recorded, and the actual conduct of the experiment itself.

Once the experimental procedure has been determined, there is another long period that the research worker spends in collecting his data. The research may require many hours in testing subjects and putting them through the experimental conditions. It is not unusual that this stage of the research may require anywhere from several months to a year to complete.

After the data have been collected, there is still the problem of analyzing the data. Even with modern calculating machines, many hours may be required before the statistical analysis is complete. With the completion of the statistical analysis, the good research worker will still spend a good many additional hours on the interpretation of his results and in speculation about new problems which his results raise.

At each of the various stages of a research project the research worker encounters problems that are unique to his particular research, problems that require solutions and decisions before he can move on to the next stage. It is in the doing of research, in the finding of solutions, in the making of decisions, that one learns to appreciate the complexities and difficulties of research, as well as the satisfactions that come from having planned and executed a well-designed experiment. This is a difficult feeling to convey in words, but those who have experienced the feeling of accomplishment which comes from the completion of even a minor but successful research problem will recognize what is meant.

Writing a good research report and making the results of the research available to other investigators is itself an arduous task. The article must be in keeping with scientific brevity, yet it should be sufficiently complete so that the reader will understand in detail the experimental procedure as well as the major findings. The article which eventually appears in one of the psychological journals may take only ten or fifteen minutes to read, but it is a tribute to the patience, persistence, and ingenuity of the research worker.

References

Allport, G. W. *Personality.* New York: Holt, 1937.

Attneave, F. A method of graded dichotomies for the scaling of judgments. *Psychol. Rev.,* 1949, *56,* 334-340.

Bartlett, M. S. Some examples of statistical methods of research in agriculture and applied biology. *J. R. statist. Soc. Suppl.,* 1937, *4,* 137-170.

Bugelski, B. R. A note on Grant's discussion of the Latin square principle in the design and analysis of psychological experiments. *Psychol. Bull.,* 1949, *46,* 49-50.

Cantril, H (Ed.) *Gauging public opinion.* Princeton: Princeton Univ. Press, 1944.

Chapin, F. S. *Experimental designs in sociological research.* New York: Harper, 1947.

Christiansen, H. F. Reported in F. S. Chapin, *Experimental designs in sociological research.* New York: Harper, 1947. Pp. 99-109.

Cochran, W. G., & Cox, Gertrude M. *Experimental designs.* New York: Wiley, 1950.

Cronbach, L. J. *Essentials of psychological testing.* New York: Harper, 1949.

Dixon, W. J., & Massey, F. J., Jr. *Introduction to statistical analysis.* New York: McGraw-Hill, 1951.

Edwards, A. L. *Statistical analysis.* New York: Rinehart, 1946.

Edwards, A. L. Homogeneity of variance and the Latin square design. *Psychol. Bull.,* 1950a, *47,* 118-129.

Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950b.

Edwards, A. L. Statistical theory and research design. In C. P. Stone & D. W. Taylor, (Eds.), *Annual review of psychology.* Vol. II. Stanford, Cal.: Annual Reviews Inc., 1951a. Pp. 335-352.

Edwards, A. L. Balanced Latin square designs in psychological research. *Amer. J. Psychol.,* 1951b, *64,* 598-603.

Edwards, A. L., & Cronbach, L. J. Experimental design for research in psychotherapy. *J. clin. Psychol.,* 1952, *8,* 51-59.

Edwards, A. L., & Kenney, Kathryn C. A comparison of the Thurstone and Likert techniques of attitude scale construction. *J. appl. Psychol.,* 1946, *30,* 72-83.

Edwards, A. L., & Kilpatrick, F. P. A technique for the construction of attitude scales. *J. appl. Psychol.,* 1948, *32,* 374-384.

Edwards, A. L., & Thurstone, L. L. An internal consistency check for scale values determined by the method of successive intervals. *Psychometrika,* 1952, *17,* 169-180.

Fisher, R. A. *The design of experiments.* (3rd ed.) Edinburgh: Oliver and Boyd, 1942.

Fisher, R. A., & Yates, F. *Statistical tables for biological, agricultural, and medical research.* (3rd ed.) New York: Hafner, 1949.

Freedman, R. Incomplete matching in ex post facto studies. *Amer. J. Sociol.,* 1950, *55,* 485-487.

Freedman, R., & Hawley, A. H. Unemployment and migration in the depression (1930-1935). *J. Amer. statist. Ass.,* 1949, *44,* 260-272.

Goodenough, Florence L. *Mental testing.* New York: Rinehart, 1949.

Goulden, C. H. *Methods of statistical analysis.* New York: Wiley, 1939.

Grant, D. A. The Latin square principle in the design and analysis of psychological experiments. *Psychol. Bull.,* 1948, *45,* 427-442.

Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.

Guilford, J. P. The computation of psychological values from judgments in absolute categories. *J. exp. Psychol.,* 1938, *22,* 32-42.

Gulliksen, H. *Theory of mental tests.* New York: Wiley, 1950.

Guthrie, E. R. *The psychology of learning.* (Rev. ed.) New York: Harper, 1952.

Guttman, L. The Cornell technique for scale and intensity analysis. *Educ. psychol. Measmt.,* 1947, *7,* 247-280.

Harris, M., Horvitz, D. G., & Mood, A. M. On the determination of sample sizes in designing experiments. *J. Amer. statist. Ass.,* 1948, *43,* 391-402.

Hevner, Kate An empirical study of three psychophysical methods. *J. gen. Psychol.,* 1930, *4,* 191-212.

Hoel, P. G. *Introduction to mathematical statistics.* New York: Wiley, 1947.

Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. *Experiments on mass communication.* Vol. III. *Studies in social psychology in World War II.* Princeton: Princeton Univ. Press, 1949.

Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.

Irwin, F. W., Armitt, Fannie M., & Simon, C. W. Studies in object-preference. I. The effect of temporal proximity. *J. exp. Psychol.,* 1943, *33,* 64-72.

JOHNSON, P. O. *Statistical methods in research.* New York: Prentice-Hall, 1949.

JONES, L. V. Tests of hypotheses: one-sided vs. two-sided alternatives. *Psychol. Bull.,* 1952, *49,* 43-46.

KEMPTHORNE, O. *The design and analysis of experiments.* New York: Wiley, 1952.

KENDALL, M. G. *Rank correlation methods.* London: Griffin, 1948.

KOFFKA, K. *Principles of Gestalt psychology.* New York: Harcourt, Brace, 1935.

KURTZ, A. K., JASPEN, N., & ASH, P. An efficient method of partitioning populations among control and experimental groups. *Educ. psychol. Measmt.,* 1951, *11,* 578-586.

LEWIN, K. *A dynamic theory of personality.* New York: McGraw-Hill, 1935.

LEWIN, K. *Principles of topological psychology.* New York: McGraw-Hill, 1936.

LIKERT, R. A technique for the measurement of attitudes. *Arch. Psychol.,* 1932, No. 140.

LINDQUIST, E. F. *Statistical analysis in educational research.* Boston: Houghton Mifflin, 1940.

MCNEMAR, Q. *Psychological statistics.* New York: Wiley, 1949.

MCNEMAR, Q. On the use of Latin squares in psychology. *Psychol. Bull.,* 1951, *48,* 398-401.

MARKS, M. R. Two kinds of experiment distinguished in terms of statistical operations. *Psychol. Rev.,* 1951, *58,* 179-184.

MATHER, K. *Statistical analysis in biology.* (2nd ed.) New York: Interscience, 1947.

MILLER, N. E., & DOLLARD, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.

MOOD, A. M. *Introduction to the theory of statistics.* New York: McGraw-Hill, 1950.

MURPHY, G., MURPHY, LOIS B., & NEWCOMB, T. M. *Experimental social psychology.* New York: Harper, 1937.

MURRAY, H. A. (Ed.) *Explorations in personality.* New York: Oxford Univ. Press, 1938.

SAFFIR, M. A. A comparative study of scales constructed by three psychophysical methods. *Psychometrika,* 1937, *2,* 179-198.

SEARS, R. R. Survey of objective studies of psychoanalytic concepts. New York: *Soc. Sci. Res. Coun. Bull.,* 1943, No. 51.

SKINNER, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.

SNEDECOR, G. W. *Statistical methods.* (4th ed.) Ames, Iowa: State College Press, 1946.

SOLOMON, R. L. An extension of control group design. *Psychol. Bull.,* 1949, *46,* 137-150.

STOUFFER, S. A. Some observations on study design. *Amer. J. Sociol.,* 1950, *55,* 355-361.

THORNDIKE, R. L. Regression fallacies in the matched groups experiment. *Psychometrika,* 1942, *7,* 85-102.

THURSTONE, L. L., & CHAVE, E. J. *The measurement of attitude.* Chicago: Univ. Chicago Press, 1929.

TIPPETT, L. H. C. *The methods of statistics.* (3rd ed.) London: Williams and Norgate, 1941.

TOLMAN, E. C. *Purposive behavior in animals and men.* New York: Appleton-Century, 1932.

WALKER, HELEN M. *Elementary statistical methods.* New York: Holt, 1943.

WOODWORTH, R. S. *Experimental psychology.* New York: Holt, 1938.

WRIGHT, H. F. The influence of barriers upon strength of motivation. *Contr. psychol. Theor.,* Durham, N. C.: Duke Univ. Press, 1937.

CHAPTER 8

Selected Quantitative Techniques

FREDERICK MOSTELLER AND ROBERT R. BUSH*
Harvard University

In the preceding chapter, Edwards has given a general introduction to the problem of the design and execution of experiments. That chapter included a general background of commonly used statistical methods. In this chapter we follow with certain additional quantitative techniques that should be of value in the design of experiments or the analysis of their results. Three main considerations have guided us in choosing this set of techniques. First, we have for the most part avoided presenting again materials available in the several excellent elementary texts on statistics widely available to psychologists; indeed, we assume our readers have already digested one such text. We must emphasize therefore that the present chapter is not designed for the beginner but is an advanced chapter for a person concerned with a particular research problem or with a general interest in the analysis of data. It is ironic that an advanced chapter contains some of the very easiest techniques to apply. Second, we have tried to present topics that students and colleagues frequently encounter and inquire about, but for which it is not usually easy to provide a reference. Third, we have tried to deal with some problems which are not so often raised, but which we believe deserve much more attention than they ordinarily get. Wherever possible, we have tried to make the discussion self-contained, but it will be understood that some extensive tables and expositions could not be reproduced here. On the other hand, except for common tables of the normal, t, χ^2, F, squares and logarithms, and the not-so-common inverse sine tables, we have provided necessary tables or charts to accompany any method we recommend.

The material in this chapter, as the title implies, represents a rather arbitrary selection of techniques. Therefore, the reader should not expect much continuity between the main sections. But all the techniques have one common bond: each has been useful in helping social psychologists solve some of their quantitative problems.

Instead of a simple table of contents, we give the titles of the sections, together with a few remarks about the techniques discussed.

1. *Operating Characteristics or Power Curves for the Common Statistical Tests of Significance.* How large a sample do I need? Can I detect a real difference between experimental conditions? The purpose of this section is to discuss such questions, insofar as they can be discussed in the light of considerations preliminary to the experiment. This section is more difficult to read than the remainder of the chapter.

---

*We are indebted to the Laboratory of Social Relations for facilitating this work. Doris Entwisle assisted us in all but the final stages of preparing the manuscript, Cleo Youtz assisted in preparing the manuscript for publication, and Lotte Bailyn, Dwight Bartlett, and Thurlow Wilson assisted in preparing certain tables. We are also indebted to various persons and organizations for permission to reproduce tables or materials. Any errors are our responsibility, of course. The material on operating characteristics is largely drawn from Ferris, Grubbs, and Weaver (1946); permission for reproduction of the figures was given by these authors and the Institute of Mathematical Statistics. Tables for the short-cut method of allowances were made available by R. F. Link and D. L. Wallace, who produced them in connection with research sponsored by the Office of Naval Research. We also had advice from J. W. Tukey on the allowances section. L. E. Moses, the editors of the *Psychological Bulletin*, and the American Psychological Association gave permission to draw heavily on Moses (1952) for the section on nonparametric methods. Permission to reprint Dixon and Mood's (1946, pp. 560-562) sign test tables (VI and VII) was given by those authors and the American Statistical Association. Table VIII for Wilcoxon's (1945) signed rank test is reproduced by permission of its author and the Editor of *Biometrics*. Kruskal and Wallis and the American Statistical Association gave permission to reproduce the significance levels for the ranked one-way analysis of variance (Table XIII). Tables XIV and XV are derived from more extensive tables by Swed and Eisenhart (1943).

2. *Short-cut Allowances in the Analysis of Variance.* When a number of groups of subjects are studied under differing experimental or field conditions, it is desirable to be able to recognize differences between specific pairs of group outcomes. The classical analysis of variance does not do this. A very easy method for making such comparisons is provided.

3. *The Matching Problem.* One extremely common form of experimental design in social psychology involves subjects or judges who try to place pictures, statements, protocols, or other similar stimuli into categories. These subjects are given scores equal to the number of times they place objects in the correct category. Or pairs of judges get scores equal to the number of objects they place in the same category—for example, in reliability studies. Although this technique is frequently used, students are usually at a loss to compare the data with chance outcomes because the appropriate discussion does not usually appear in an elementary text.

4. *Nonparametric Methods.* These techniques are designed to replace the usual statistical tests in order to avoid stringent assumptions. Usually nonparametric methods are easy to apply, and they have an intuitive appeal for many research workers.

4a. *The Sign Test.* When matched pairs of subjects are studied under two conditions, the better condition is the one scoring higher in most pairs. A table of sample sizes necessary to detect given differences is provided.

4b. *The Median Test.* A simple method for comparing two samples. It is one replacement for the *t*-test in unmatched groups.

4c. *Wilcoxon's Matched Pairs Test Using Signed Ranks.* Like the sign test and the median test, this is a replacement for the *t*-test, and it seems to be the best nonparametric replacement known for matched pairs.

4d. *Mann-Whitney Test.* Ranking methods are used for testing differences between two unmatched groups.

4e. *Whitney's Extension of the U-Statistic.* A device for testing whether one of three groups performed best or, alternatively, a device for testing whether performance of the three groups was in a predicted order.

4f. *The Problem of m Rankings.* Do judges generally agree about the ranking of a set of objects? A method is given for testing the concordance of several rankings.

4g. *The Complete Ranked One-Way Analysis of Variance.* The usual analysis of variance is replaced by one using ranks. This is the natural generalization of the Mann-Whitney test to more than two groups of subjects.

4h. *Wald-Wolfowitz Run Test.* Do high and low observations alternate too frequently? Or do certain kinds of observations seem to be bunched together in time, space, or size instead of being evenly spread? The run test is a method of studying over- or underclustering of sequences of observations.

5. *Quick Methods.* Devices for using the range as a measure of variation in small samples for those who are weary of squaring numbers, but are willing to lose a little information in the analysis of data.

6. *Transformations.* Data that arise from counting things (successes, acts, choices, etc.) sometimes contain extra information about variability that can be obtained by adjusting or transforming the original observations. If there are important uncontrolled variables in the study, analysis of transformed data may reveal this fact.

7. *Combining Tests of Significance.* On a given psychological problem some studies have shown significant results, while others have not. How can we combine all these results to make one grand summary? Pitfalls and pros and cons of three techniques are described.

Operating Characteristics or Power Curves for The Common Statistical Tests of Significance

A question of general interest to an investigator planning an experiment is whether his design, including sample size, has a good chance of detecting differences among conditions, methods, or ethnic, age, or other population groups. The primary value of a study of group differences is lost if the investigation is so insensitive that important differences are obscured by sampling variation. We seldom enter an investigation completely in the dark about the magnitudes of differences we expect to observe (although there seems to be great cultural value in pretending no one has ever done anything like our experiment before). This general information about magnitudes of effects and the variation ordinarily observed can be very useful in planning a study. Of course, one does not know exactly what to expect, but it is obvious, for example, that a person comparing groups on a 6-point rating scale is unlikely to find average group differences as large as 4 points. Many paper and pencil tests have been standardized on known populations, so some notion of the means and standard deviations is available for them. Similarly, various experimental devices have been used in many different situations and for them various kinds of baseline data are available to the investigator willing to take the trouble to look. These data can be used in a formal way in the design of an experiment to see what size of difference the experiment is likely to detect, or to help in revising the experiment if it is too insensitive to possible differences. To illustrate the use of such information it will be necessary for us to go into more detail on the matter of Type II errors than did Edwards at the close of his section on *Tests of Significance* in Chapter 7. The reader may wish to review that section before proceeding.

If we state a null hypothesis, draw a sample, and make a statistical test such as the t-test or the F-test, we may make a mistake in our final decision. These mistakes or errors have been classified by Neyman and Pearson as:

Type I error: rejecting the null hypothesis when it is true.

Type II error: accepting the null hypothesis when it is false.

The chosen level of significance, for example, 0.05, controls the probability of making a type I error, and this probability does not change. This is the error that has customarily received the lion's share of experimenters' attention. The probability of making a type II error is not constant, rather it depends on the wrongness of the null hypothesis, the sample size, the significance level chosen, and the particular method of testing significance. Everyone agrees that it is important in experimental or survey design to give consideration to the type II error. The experimenter wants to be able to detect a difference when there is a fairly substantial difference — otherwise he would not be bothering to do experiments. In the social sciences attention to type II errors has been scanty, although people quite often approach the problem obliquely by wondering how many observations they should take. One reason for this neglect of the type II error may be that materials for studying the type II error are not readily available. However, Ferris, Grubbs, and Weaver (1946) present curves showing the probability of making type II errors for some common statistical tests. These curves are called "operating characteristics." The power of a test is the probability of *not* making a type II error.

Strictly speaking, the operating characteristic for a test gives the probability of accepting the null hypothesis for each possible situation, and the power curve gives the probability of rejecting the null hypothesis for each possible situation. When the definition is given in terms of type II errors a slight confusion can arise about the existence of a type II error when the null hypothesis is, in fact, true. Nevertheless, it seems worth while to emphasize the notion of type II errors in the text.

The term "operating characteristic" is such a good one that it is used for more than one thing. In Chapter 9, the term is used to refer to the behavior of an item on a test. The similarity in idea is not as remote as it would seem at first glance.

It follows that the power of a test and its operating characteristic are two ways of giving the same information. There are cases, particularly for nonparametric tests, where it is not easy (in fact, it is often impossible) to write down the power of the test. The reason is that nonparametric tests offer the possibility of testing a great variety of hypotheses all at once. We shall see later that in order to know how strong or powerful a test is we need to know the specific distribution from which the sample is drawn, exactly what the null hypothesis is and, even more, exactly what the alternative hypothesis is. This section will discuss cases where we do know distributions and hypotheses explicitly.

We begin by describing the operating characteristic for a binomial situation, where computation is easy. Suppose we have a factory making "fair" coins, coins that have a probability of $\frac{1}{2}$ of coming up heads each time they are spun. The inspection procedure for each coin consists in spinning the coin 5 times and observing the outcomes. The null hypothesis is that the coin has a probability $\frac{1}{2}$ of coming up heads. The alternative hypothesis is that the probability is not $\frac{1}{2}$. Coins with better than a 50-50 chance of coming up heads will produce more heads than fair coins. Our acceptance and rejection procedure is taken arbitrarily to be

(1) accept the coin as "fair" if 1, 2, 3, or 4 of the 5 coins fall heads,

(2) reject the coin as "unfair" if 0 heads or 5 heads appear.

In the language of testing hypothesis, the null hypothesis H_0 and the alternative hypothesis H_1 are

$$H_0 : p = \tfrac{1}{2},$$

$$H_1 : p \neq \tfrac{1}{2},$$

where p is the probability of a head on a single spin.

The significance level α is the sum of the probabilities associated with rejection when the coin is, in fact, fair. Since the probability of 5 heads is $\frac{1}{32} = (\frac{1}{2}) \cdot (\frac{1}{2}) \cdot (\frac{1}{2}) \cdot (\frac{1}{2}) \cdot (\frac{1}{2})$, and that for 0 heads is similarly $\frac{1}{32}$, the value of $\alpha = \frac{1}{16} = 0.0625$. This is also the probability of a type I error. The probability of not making a type I error is $1 - \alpha = 0.9375$.

When the actual value of p is different from 0.50, the probability of rejection is $p^5 + (1 - p)^5$, the two terms corresponding to the probabilities of 5 heads or 5 tails in 5 throws. The significance level, of course, remains 0.0625. In fact, the formula just given represents the power of the test, for by letting p take on values different from ½ we can compute the probabilities of rejecting the null hypothesis for every value of p. For the operating characteristic, we have

$$\text{probability of accepting } H_0 = 1 - p^5 - (1 - p)^5.$$

A table of this probability of accepting H_0 against p is the operating characteristic. This is given under Test 1 in Table I. We see that when the coin has a bias of 0.1 in favor of heads or tails (i.e., $p = 0.6$ or 0.4), the probability of accepting it as fair is still 0.91, scarcely different from the 0.9375 that goes with fair coins. We accept coins with $p = 0.8$ (these coins come up heads 4 times out of 5 on the average) about two-thirds of the time. How bad must coins be before we will be almost certain (say 95% sure) of rejecting them with our test? The coin has to be so defective that it throws heads (tails) 99 times out of 100. What has the study of this operating characteristic told us? If we are the "fair coin" factory owner we would feel that such a test would scarcely be worth using because we would not be able to give much of a guarantee about single coins. We have discovered that the particular test is rather ineffective for detecting differences from the null hypothesis $p = ½$. How can it be improved? We could change the significance level. We don't like to increase it because that would increase the probability of type I errors as well as decrease the probability of type II errors. Increasing the significance level lowers the entire curve but we have to pay for this by scrapping more perfect coins. We are already scrapping about 7% of the perfect coins if there are any—why scrap more? We can't really find a more sensitive statistical test based on the same number of observations. (We could, of course, take dimensional measurements—these would take us away from this introductory discussion—but we would again wish to study the operating characteristic of new test statistics and similar problems would crop up.) The only other solution, if we wish to improve the inspection procedure, is to increase the number of trials per coin and adjust the rejection numbers.

TABLE I

OPERATING CHARACTERISTIC FOR BINOMIAL TESTS OF $p = 0.5$.

In Test 1, the sample size is 5, and we reject if 0 or 5 successes appear.

In Test 2, the sample size is 19, and we reject if 0-5 or 14-19 successes appear.

True p	*Test 1* Significance level 0.0625 P (accepting $p = 0.5$)	*Test 2* Significance level 0.0636 P (accepting $p = 0.5$)
1.0	0.0000	0.0000
0.9	0.4095	0.0086
0.8	0.6720	0.1631
0.7	0.8295	0.5260
0.6	0.9120	0.8340
0.5	0.9375	0.9364
0.4	0.9120	0.8340
0.3	0.8295	0.5260
0.2	0.6720	0.1631
0.1	0.4095	0.0086
0.0	0.0000	0.0000

Before discussing this point, let us relate this same operating characteristic to a social science investigation. Instead of "fair coin" producers, we produce some social psychology hypotheses for test. Scientists have fertile minds and produce hypotheses in quantity and, of course, there are many of them producing these ideas. Many of the hypotheses are of a tendential sort —groups of individuals of type A differ from those of type B on such and such a measure. Suppose for each such hypothesis we get 5 matched pairs of individuals, and whenever all 5 type A members score higher (lower) we reject the hypothesis of equality of the two kinds of individuals on the measure. The operating characteristic for Test 1 describes the performance of this new test. The hypothesis that a difference exists corresponds to the alternative hypothesis in statistical language: $p \neq ½$. The null hypothesis is that no difference exists: $p = ½$. Now only half the hypotheses in which 87% of the type A individuals in the population have larger values on the measure than their type B mates will be "accepted" or confirmed on the basis of this test. The shape and properties of this particular operating curve turn statisticians prematurely gray when a small pilot study is proposed to see if there is "anything in" a particular hypothesis. It may be a wonderful hypothesis, but we are not likely to get much support for it from this test. However, a small pilot study may be very useful from the standpoint of planning and design. An intelligent examination of an operating characteristic before and during the design of an experiment can suggest not only the order of magnitude of the number of observations that may be required, but it can also suggest that some experiments will almost certainly yield equivocal results because of lack of facilities, or inadequate measuring instruments, or lack of time, and therefore that the experiment should either not be undertaken or its plan revised.

What would it be worth in the social-science-hypothesis experiment to use a sample of about 20? (We use 19 so that we can make the significance levels of this test and the test with 5 observations almost identical.) The test now consists of rejecting the null hypothesis if 0, 1, 2, 3, 4, 5, or 14, 15, 16, 17, 18, or 19 outcomes of the same kind are observed. The operating characteristic is that of Test 2 in Table I. With $n = 5$ and $p = 0.8$, the probability of accepting the null hypothesis was about 0.67, but with $n = 19$, the probability of accepting the null hypothesis is only 0.16. How big a p do we need to be almost certain (0.95) to reject the null hypothesis? A p of about 0.85 will be rejected this often. With $n = 5$, a p of 0.99 was required. So the increase from 5 observations to 19 observations does make a substantial difference in ability to discriminate between hypotheses, as Table I shows.

What about a man who says no null hypothesis is ever correct? If we had 200,000 flips of a coin, for example, almost certainly we would discover that it or the method of spinning was imperfect. The point seems to us to go something like this. It is unlikely that any mathematical model for a situation is perfect, and therefore we agree that it is unlikely that any null hypothesis is really true. We are interested in detecting departures from the model, and therefore from the null hypothesis, that are "substantial," i.e., that are consequential as far as decisions are concerned. Reasonable size samples will not discriminate well between measures that are very close, say $p = 0.500$ and $p = 0.501$. Such close results are not of great consequence in most behavioral situations. Consequently, from the investigator's viewpoint, there is really a "don't care" band in the neighborhood of the value of the null hypothesis, and as we depart further and further from the null hypothesis value, the investigator takes a more interested attitude toward the difference. The operating characteristic behaves in an analogous manner. In the vicinity of the null hypothesis value, the null hypothesis is rarely rejected, and values nearby are not often rejected either. As the departure from the null hypothesis increases, an operating characteristic shows that we are more and more likely to detect the departure. We do not wish to confuse the issues of practical importance with "statistical significance." On the other hand, when an investigator chooses a test, a significance level, and a sample size, he has almost automatically implied that his attitude toward differences (operationally) is consistent with the shape of the operating characteristic.

Suppose the investigator says: "I don't like making all these mistakes—I want to reject for sure if the true value is outside 0.45 to 0.55 and forget about the matter otherwise." This statement specifies the operating characteristic precisely and, incidentally, the sample size is unwittingly determined too. The sample size is infinite, so our investigator will not get enough data in his lifetime to settle the matter as far as he is concerned. Once this point is firmly established in an investigator's mind, an interest in operating characteristics is usually apparent. For further discussion, see *Number of Subjects,* in Chapter 7.

For purposes of illustration we have discussed a binomial example. We can apply the same ideas to the usual statistical tests, and we shall explain how this is done, but we shall not derive the operating characteristics. Instead, we provide charts for some commonly used tests, together with illustrations of their use. In the operating characteristic figures the significance level chosen is 0.05. The appropriate tables: χ^2, F, t, and the normal, will be found in many elementary texts.

We shall discuss two applications of operaing characteristics to chi-square tests. These chi-square tests are used to compare observed sampling variation with the variation expected from a standard population. Thus we want to test whether the scores of our sample are more homogeneous or more heterogeneous than those of a large known population tested on some other occasion. (There may be confusion when we call these "chi-square tests" because social scientists usually employ this name exclusively for contingency tables or for goodness-of-fit tests. These latter tests are not being discussed.)

Power function of the chi-square test. The present chi-square test is used to accept or reject the hypothesis that the true standard deviation

σ_1 of a normal population equals a specified value σ.

The hypotheses are

$$H_0:\ \sigma_1 = \sigma,$$

$$H_1:\ \sigma_1 = \lambda\sigma,\ \ (\lambda > 1)\,.$$

Thus, we are making a one-sided test (for the distinction between one- and two-sided tests, see *Tests of Significance,* in Chapter 7), because we are asking whether or not our standard deviation is larger than the standard (i.e., λ larger than unity). Let us suppose that we have a set of observations drawn from the normal, and that their values are $x_1, x_2, \ldots, x_n$, where n is the sample size. The test statistic is

$$\chi^2 = \frac{(n-1)\,s^2}{\sigma^2},$$

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}, \qquad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

(Hereafter if we drop the limits on the summation, it is to be understood that summations run over all the observations of the sample for which they are appropriate.)

The test consists of rejecting the null hypothesis if $\chi^2 > \chi_\alpha^2$, and accepting otherwise. Here χ_α^2 is the α percentage point of the chi-square distribution as it is commonly tabulated, i.e., large chi-square values go with small values of α. The number of degrees of freedom is one less than the number of observations.

Example of chi-square test. A standardized paper and pencil test is to be given to hospital patients. The social science hypothesis is that owing to the disease they suffer from, the patients will have more variation in their scores than will the group who served to standardize the test. The standard deviation of the test for normal groups is $\sigma = 10$. The investigator guesses that the hospitalized group will have a standard deviation of about $\sigma_1 = 20$. Thus λ is conjectured to be about $20/10 = 2$. If the value of λ is this large, and if the investigator wants to be at least 90% sure of rejecting the null hypothesis that $\sigma_1 = 10$, how big a sample does he need when using the 5% level of significance? On the horizontal axis of Fig. 1 go to $\lambda = 2.0$, then read up until a curve is found crossing the ordinate at a probability of accepting of 0.1. A sample size of 10 meets the requirement nicely. Because of the facilities available, it will be easy to get 15 observations and so the chance of rejecting is now up to 0.96, and even if λ is only 1.7 the chance is about 0.9 of rejecting the null hypothesis with $n = 15$.

Another use of the chi-square test. Fig. 2 is similar to Fig. 1. The same test statistic is used; the hypotheses are

$$H_0:\ \sigma_1 = \sigma,$$

$$H_1:\ \sigma_1 = \lambda\sigma,\ \ (\lambda < 1)\,,$$

and we reject if $\chi^2 < \chi^2_{1-\alpha}$ accept otherwise.

The test is again one-sided; we are testing whether the standard deviation of the special normal population under study is smaller than the standard value σ.

Power of the F-test. In simplest terms, the F-test is a comparison of independent sample estimates of two normal population variances. The one-sided tests inquire whether the population variance from which sample 1 is drawn is larger than the population variance from which sample 2 is drawn. We no longer have a known standard σ, as we did earlier.

The hypotheses are

$$H_0:\ \sigma_1 = \sigma_2,$$

$$H_1:\ \sigma_1 = \lambda\sigma_2,\ \ (\lambda > 1)\,.$$

The test statistic is

$$F = \frac{s_1{}^2}{s_2{}^2}\,,$$

where

$$s_1{}^2 = \frac{\Sigma(x_1 - \bar{x}_1)^2}{n_1 - 1}, \qquad s_2{}^2 = \frac{\Sigma(x_2 - \bar{x}_2)^2}{n_2 - 1}.$$

If the observed value of F exceeds the tabled value F_α, we reject the null hypothesis of equality of variances, otherwise we accept. Note that there is no question of putting the larger mean square over the smaller mean square. We are clear that σ_1 is to be larger than, or the same as σ_2; therefore we divide $s_1{}^2$ by $s_2{}^2$ to get F and enter anybody's F-table with $n_1 - 1$ and $n_2 - 1$ degrees of freedom respectively. Of course, if $s_1{}^2$ is less than $s_2{}^2$, there is no need to consult the F-table at all—we accept the null hypothesis.

Example of the F-test. We have the hypothesis that men in a certain population are more variable in their attitudes on a certain issue than are women. We measure the variability by the sample variance of a battery of questions with scores roughly normally distributed. The test statistic F is the ratio of the estimate of variance for men to the estimate of variance for women on the basis of the observed sample. If

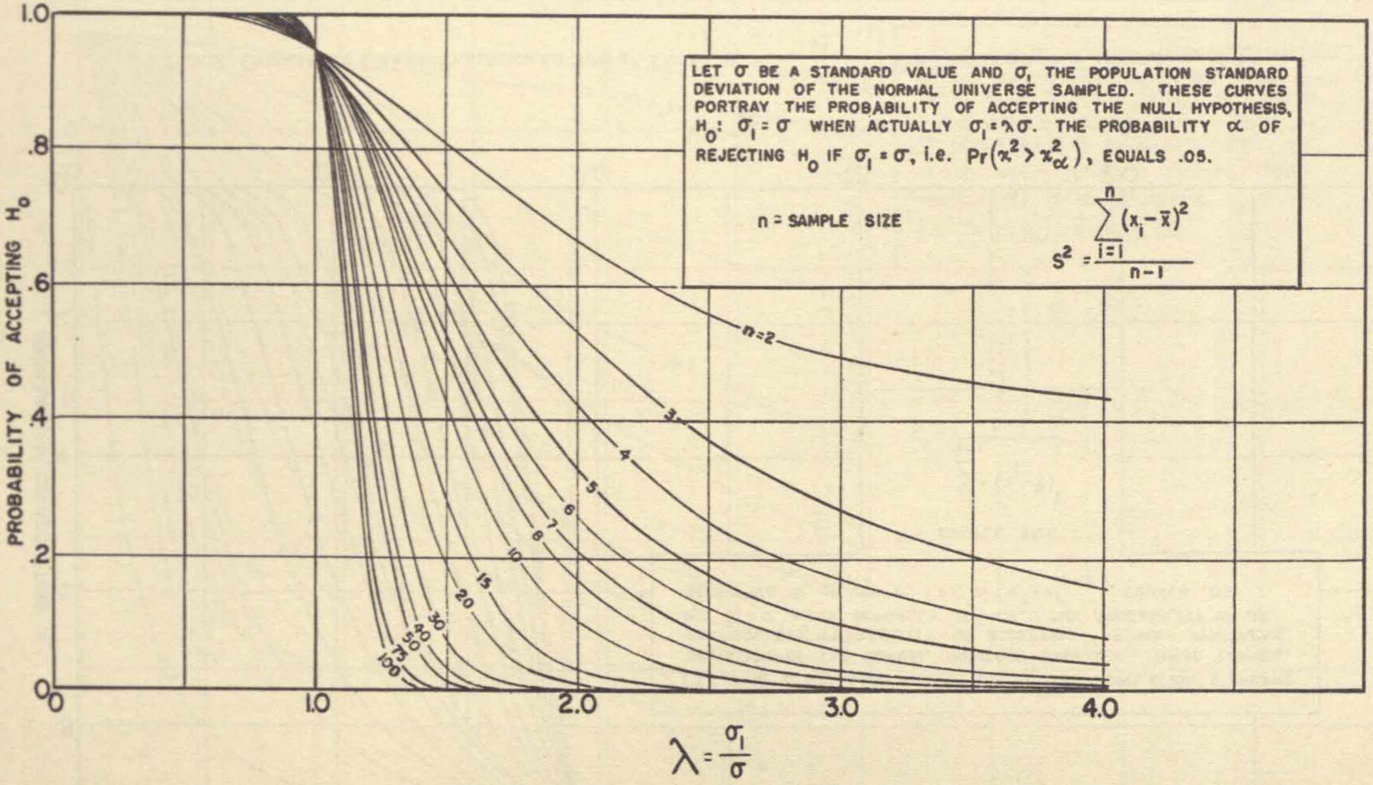


FIG. 1. OPERATING CHARACTERISTICS OF THE χ^2-TEST $\left[\chi^2 = \frac{(n-1)s^2}{\sigma^2}\right]$ FOR TESTING $\sigma_1 = \sigma$ AGAINST $\sigma_1 > \sigma$

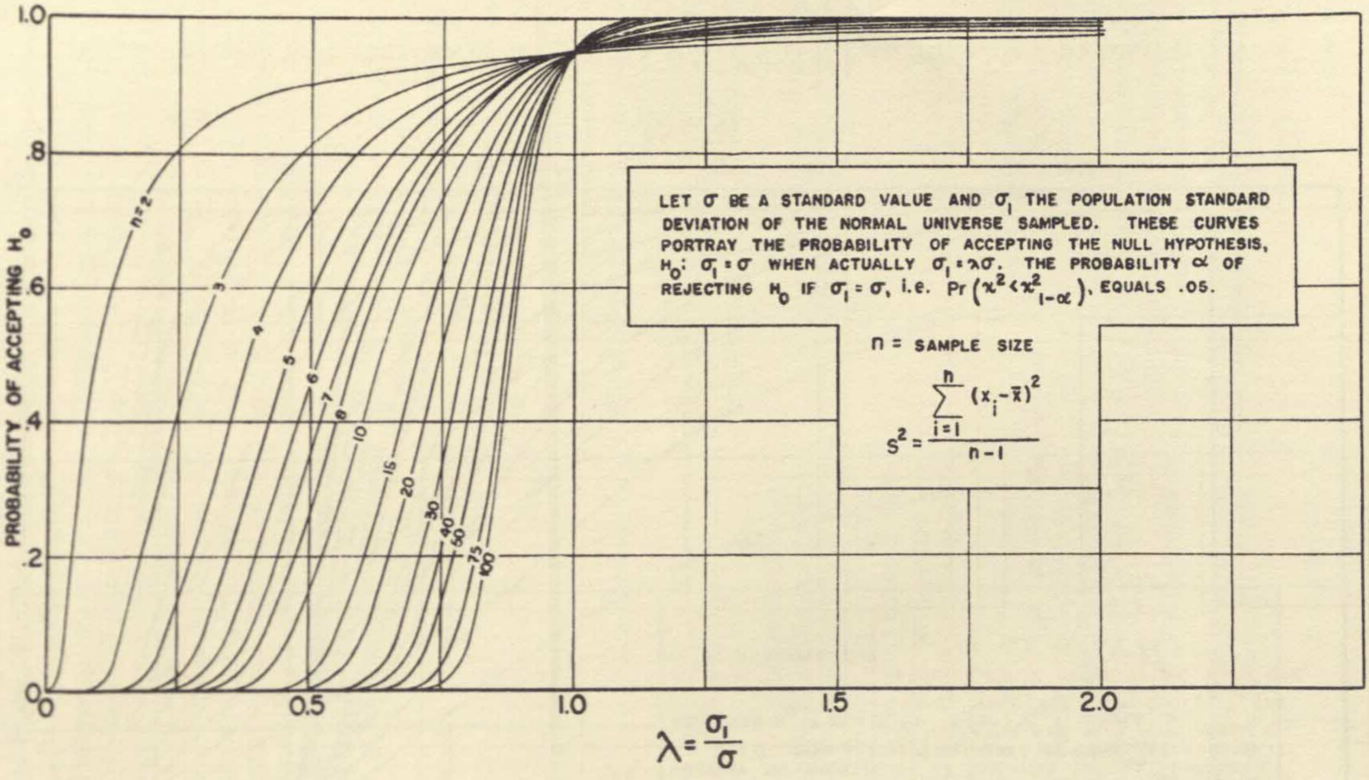
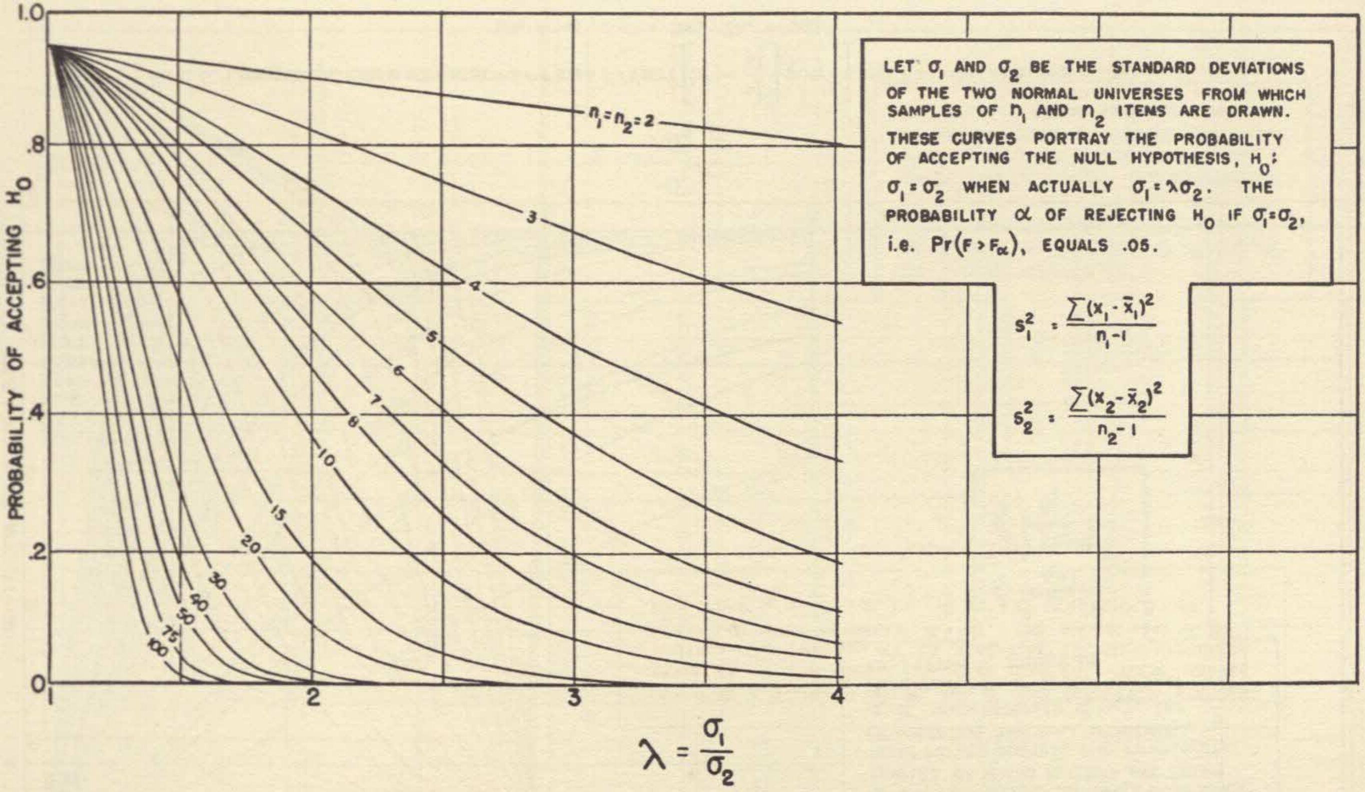


FIG. 2. OPERATING CHARACTERISTICS OF THE χ^2-TEST $\left[\chi^2 = \frac{(n-1)s^2}{\sigma^2}\right]$ FOR TESTING $\sigma_1 = \sigma$ AGAINST $\sigma_1 < \sigma$



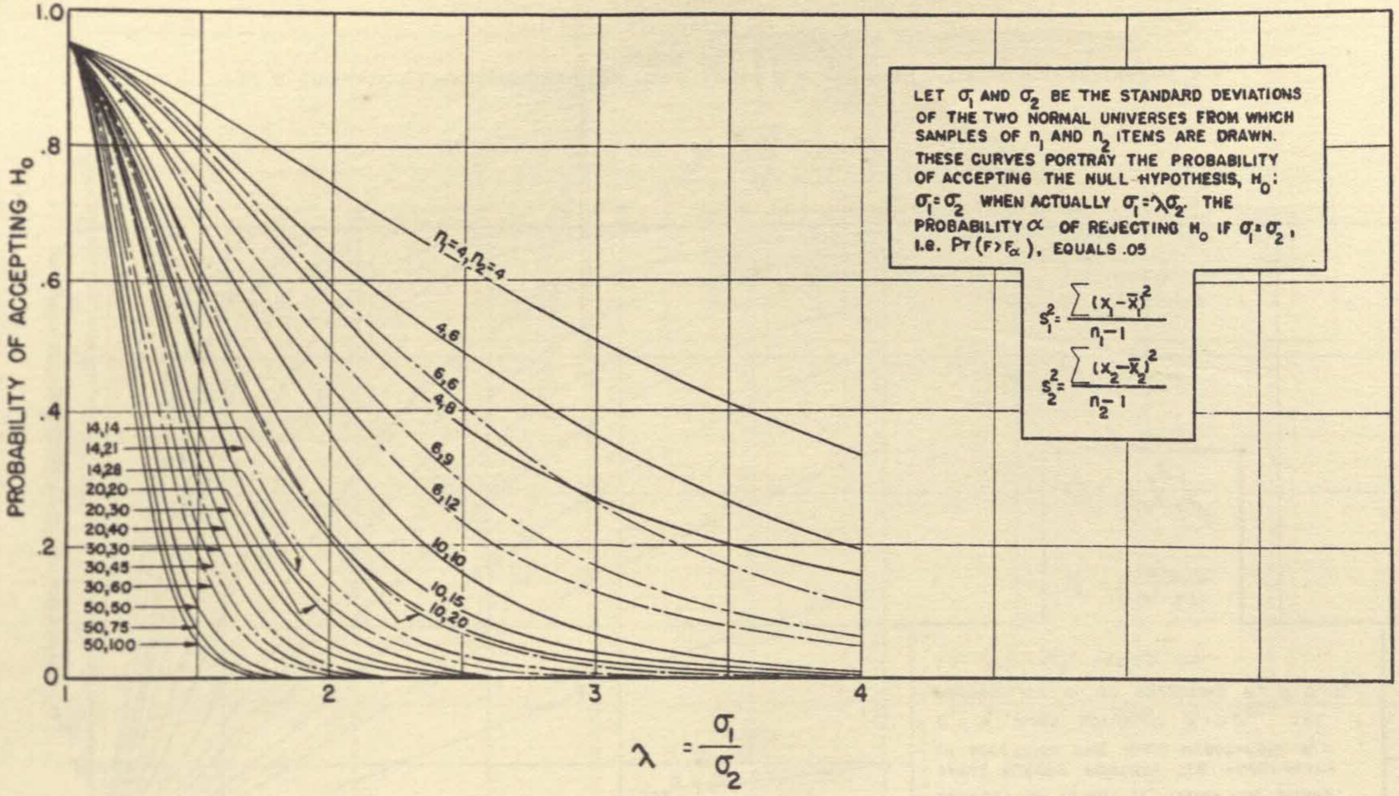
LET σ_1 AND σ_2 BE THE STANDARD DEVIATIONS OF THE TWO NORMAL UNIVERSES FROM WHICH SAMPLES OF n_1 AND n_2 ITEMS ARE DRAWN. THESE CURVES PORTRAY THE PROBABILITY OF ACCEPTING THE NULL HYPOTHESIS, H_0: $\sigma_1 = \sigma_2$ WHEN ACTUALLY $\sigma_1 = \lambda\sigma_2$. THE PROBABILITY α OF REJECTING H_0 IF $\sigma_1 = \sigma_2$, i.e. $Pr(F > F_\alpha)$, EQUALS .05.

$$s_1^2 = \frac{\sum(x_1 - \bar{x}_1)^2}{n_1 - 1}$$

$$s_2^2 = \frac{\sum(x_2 - \bar{x}_2)^2}{n_2 - 1}$$

FIG. 3. OPERATING CHARACTERISTICS OF THE F-TEST $\left[F = \frac{s_1^2}{s_2^2}\right]$ FOR TESTING $\sigma_1 = \sigma_2$ AGAINST $\sigma_1 > \sigma_2$

$(n_1 = n_2)$



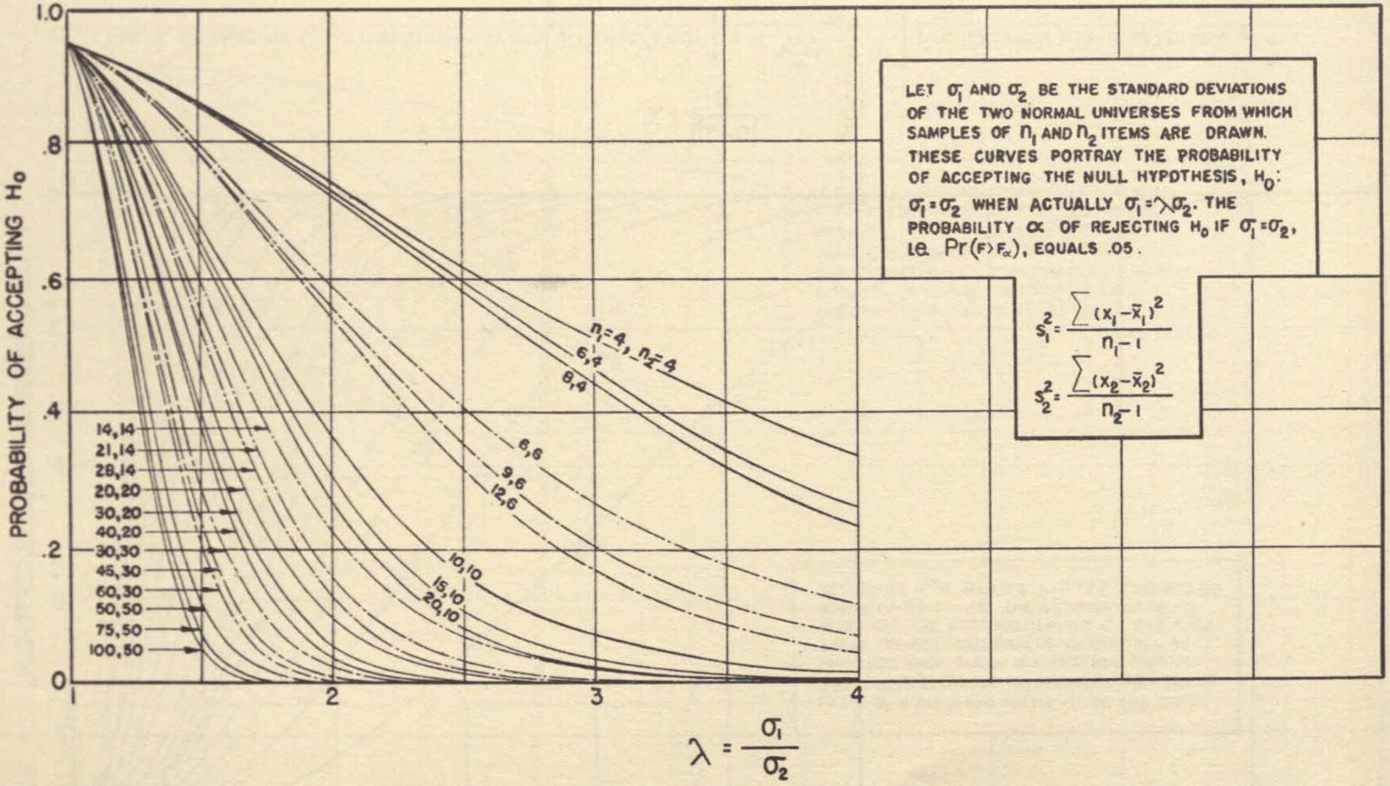
LET σ_1 AND σ_2 BE THE STANDARD DEVIATIONS OF THE TWO NORMAL UNIVERSES FROM WHICH SAMPLES OF n_1 AND n_2 ITEMS ARE DRAWN. THESE CURVES PORTRAY THE PROBABILITY OF ACCEPTING THE NULL HYPOTHESIS, H_0: $\sigma_1 = \sigma_2$ WHEN ACTUALLY $\sigma_1 = \lambda\sigma_2$. THE PROBABILITY α OF REJECTING H_0 IF $\sigma_1 = \sigma_2$, I.E. $Pr(F > F_\alpha)$, EQUALS .05

$$s_1^2 = \frac{\sum (x_1 - \bar{x}_1)^2}{n_1 - 1}$$

$$s_2^2 = \frac{\sum (x_2 - \bar{x}_2)^2}{n_2 - 1}$$

FIG. 4. OPERATING CHARACTERISTICS OF THE F-TEST $\left[F = \frac{s_1^2}{s_2^2}\right]$ FOR TESTING $\sigma_1 = \sigma_2$ AGAINST $\sigma_1 > \sigma_2$

$(n_1 = n_2\,,\ 3n_1 = 2n_2\,,\ 2n_1 = n_2)$



LET σ_1 AND σ_2 BE THE STANDARD DEVIATIONS OF THE TWO NORMAL UNIVERSES FROM WHICH SAMPLES OF n_1 AND n_2 ITEMS ARE DRAWN. THESE CURVES PORTRAY THE PROBABILITY OF ACCEPTING THE NULL HYPOTHESIS, H_0: $\sigma_1 = \sigma_2$ WHEN ACTUALLY $\sigma_1 = \lambda\sigma_2$. THE PROBABILITY α OF REJECTING H_0 IF $\sigma_1 = \sigma_2$, I.E. $Pr(F > F_\alpha)$, EQUALS .05.

$$s_1^2 = \frac{\sum (x_1 - \bar{x}_1)^2}{n_1 - 1}$$

$$s_2^2 = \frac{\sum (x_2 - \bar{x}_2)^2}{n_2 - 1}$$

Fig. 5. Operating Characteristics of the F-Test $\left[F = \frac{s_1^2}{s_2^2}\right]$ for Testing $\sigma_1 = \sigma_2$ against $\sigma_1 > \sigma_2$

($n_1 = n_2$, $2n_1 = 3n_2$, $n_1 = 2n_2$)

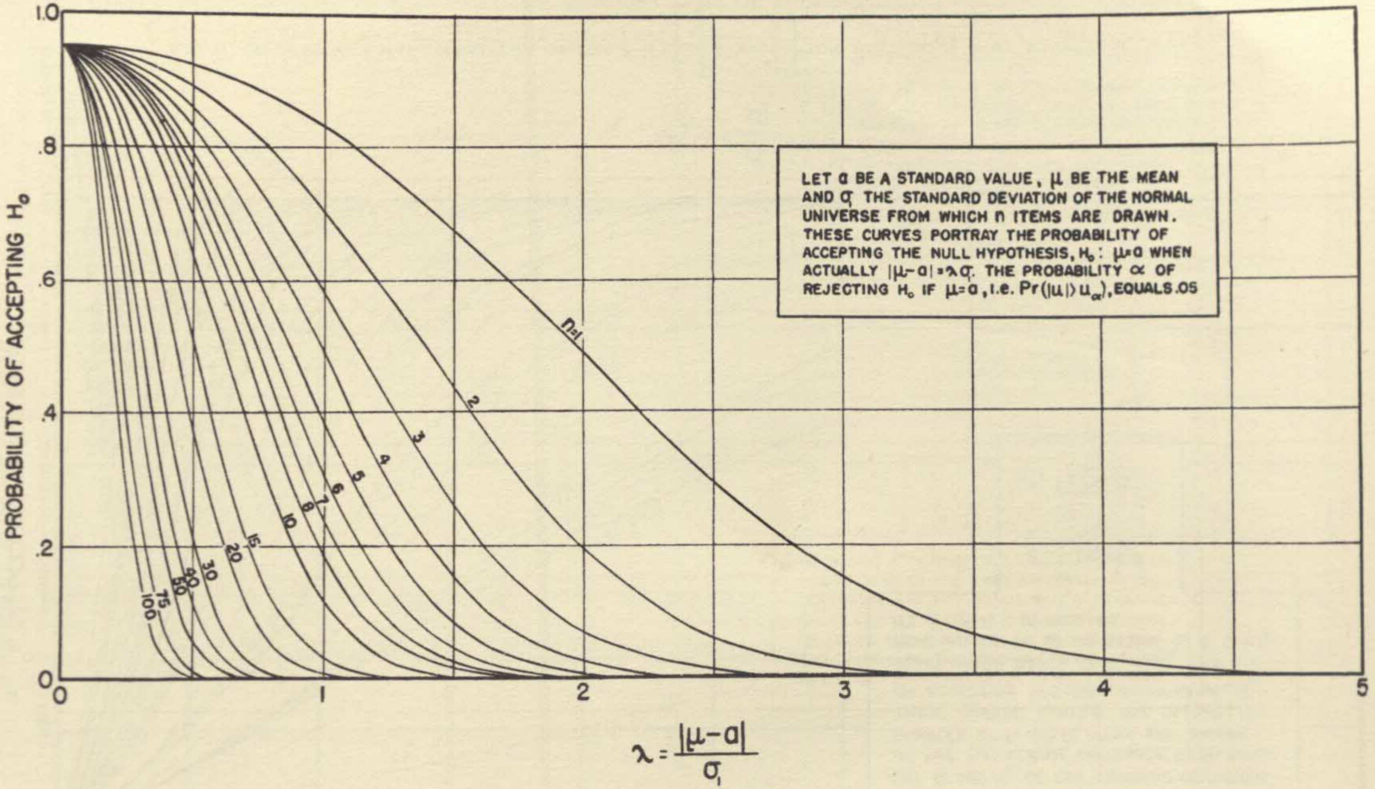


Fig. 6. Operating Characteristics of the Normal Test $\left[u = \frac{\sqrt{n}(\bar{x} - a)}{\sigma_1}\right]$ for Testing $\mu = a$ against $\mu \neq a$

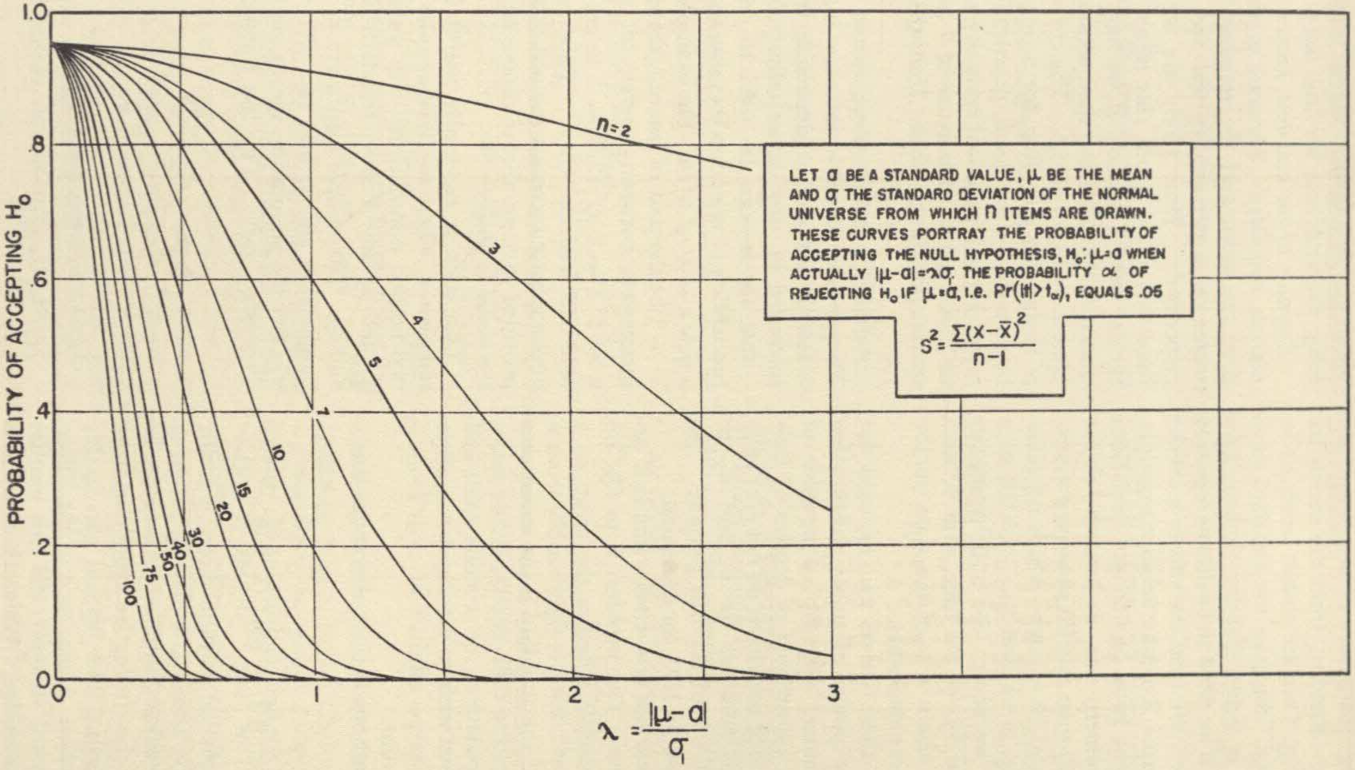


FIG. 7. OPERATING CHARACTERISTICS OF THE t-TEST $\left[t = \frac{\sqrt{n}(\bar{x} - a)}{s}\right]$ FOR TESTING $\mu = a$ AGAINST $\mu \neq a$

the sample sizes for men and for women are both 10, how much larger does the standard deviation for men have to be for us to be 80% sure of rejecting (20% sure of accepting) the null hypothesis? Reading from the curves for $n = 10$ from Fig. 3 we see that it needs to be 2.4 times as large. Suppose the standard deviation for men is twice that for women; how large a sample do we need to be sure of rejecting 90% of the time? We read the answer as about 20 in each group. To take another example, suppose we have a ratio of standard deviations of 2 (men to women) and we have the choice of a sample of 40 men and 20 women or a sample of 30 men and 30 women—which is preferable? From Fig. 5 we find that with 40 men and 20 women we accept the null hypothesis (wrongly) about 5% of the time, with 30 men and 30 women about 4% of the time, so the equal split is more desirable.

In principle, these curves can be used for the study of the power of single classification analysis of variance, with n_1 = number of groups, $n_2 - 1$ = number of degrees of freedom for error. But the charts do not give sufficiently disparate numbers for n_1 and n_2 to make them very useful for this purpose. See Ferris, *et al.* (1946, p. 188) for discussion.

Power of the normal test, known standard deviation, two-sided. The problem is to test on the basis of a sample whether the mean of a normal population departs significantly from a standard. Often the standard is the value zero (when we are testing differences), or the mean of a census population, or a well-established standard published with an intelligence or other test based on a large number of cases from a specified population.

The null hypothesis and its alternative are

$$H_0: \ \mu = a,$$
$$H_1: \ |\mu - a| = \lambda\sigma_1, \ (\lambda > 0).$$

Here μ is the true population mean, a is the standard, σ_1 is the known standard deviation of the normal population, and λ is a measure of the departure of the alternative hypothesis from the null hypothesis. The vertical bars in H_1 mean absolute value (magnitude of the difference without regard to sign) and this implies that the test is two-sided. We are just as interested whether μ is smaller than a or bigger than a. The test function is the critical ratio

$$u = \frac{\sqrt{n}\ (\bar{x} - a)}{\sigma_1},$$

where n is the number of cases in the sample, and $\bar{x}$ is the sample mean. If $|u| > 1.96$ we reject the null hypothesis at the 5% level. This is the classical critical ratio test, properly used because σ_1 is known. Figure 6 gives the operating characteristic for this test.

Example of normal test, known standard deviation, two-sided. For sixth grade school children the mean and the standard deviation of scores on a personality test are 100 and 10, respectively. How likely are we to detect a departure of 15 from this mean in a certain special class of schools if we have 50 children drawn at random from their sixth grade classes? Here $\lambda = 15/10 = 1.5$. The curve for $n = 50$ is kissing the zero line for accepting the null hypothesis, so we seem practically certain to reject. Indeed, practical certainty extends down to $\lambda = 0.7$, or a departure of 7 points in true special population mean from grand mean of 100.

Actually, not all these curves are needed. They are all related in this way: if n' and λ' are a sample size and a departure index we are interested in, then some other reference curve —that for n—can be used to determine the probability of rejection by using the new n with a new $\lambda = \lambda' \sqrt{n'/n}.$ For example, if $n' = 12$, $\lambda' = 1$ be the pair of interest, then we need not interpolate between curves, for we can use $n \doteq 3$, $\lambda = 1\sqrt{12/3} = 2$ to get $\lambda = 2$, and read 0.07 from the curve. This transformation device greatly increases the utility of the curves, provided that we always choose an n for which the curve is given.

This trick is especially useful for the two-sample test, where we have samples drawn from two normal populations with unknown means μ_1 and μ_2 and known standard deviations σ_1 and σ_2. The sample means are $\bar{x}_1$ and $\bar{x}_2$, and sample sizes are n_1 and n_2. We want to test whether the population means are equal or not (two-sided test). The test function is

$$u = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1{}^2}{n_1} + \dfrac{\sigma_2{}^2}{n_2}}}.$$

Again the hypotheses are

$$H_0: \ \mu_1 = \mu_2,$$
$$H_1: \ |\mu_1 - \mu_2| = \lambda'\sigma_1, \qquad (\lambda' > 0).$$

We take $\sigma_2/\sigma_1 = k$. Then we introduce the concept of effective sample size n'. Its formula is

$$n' = \frac{n_1 n_2}{k^2 n_1 + n_2}.$$

We use n' as if we had a single sample instead of two samples. Suppose we guess $\mu_1 - \mu_2$ is about 10, and know $\sigma_1 = 5$, $\sigma_2 = 10$; then $\lambda' = 2$ and $k = 2$. If we have sample sizes $n_1 = 5$, $n_2 = 20$, then

$$n' = \frac{5 \times 20}{4 \times 5 + 20} = 2.5.$$

Sadly enough, there is no curve for $n' = 2.5$, but we can use the transformation trick. If we take $n = 1$, then

$$\lambda = \lambda' \sqrt{\frac{n'}{n}} = 2 \sqrt{\frac{2.5}{1}} = 3.16.$$

This value of λ gives a reading for the curve $n = 1$ of about 0.11 as the probability of accepting.

Power of the t-test, two-sided. This is the classic test used to accept or reject the hypothsis that the true mean μ of a normal distribution equals a standard a.

The hypotheses are

$$H_0: \ \mu = a,$$
$$H_1: \ | \mu - a | = \lambda\sigma_1, \qquad (\lambda > 0).$$

The test function is

$$t = \frac{\sqrt{n}\,(\bar{x} - a)}{s}, \quad s^2 = \frac{\Sigma\,(x_i - \bar{x})^2}{n - 1}.$$

We reject the null hypothesis if $t > t_\alpha$, where t_α is the two-sided percentage point of the ordinary Student t-table (example: $t_{0.05} = 2.086$ with 20 d.f., i.e., $n = 21$).

Example of the t-test, two-sided. Edwards illustrates the t-test with a comparison of performance of matched pairs working under conditions of noise, where A members have the attitude that noise increases output, while B members have the contrary attitude (Edwards, 1946, p. 173). Let us take his observed measures of mean and errors as true and ask how large a sample of matched pairs we need to be 90% sure of getting significance. He has

$$\sigma_1 = 0.63, \quad | \mu - a | = 1.$$

Then $\lambda = 1/0.63 = 1.6$, and from Fig. 7 we see that we need about 6 matched pairs.

Power of the t-test, two-sided, for testing difference between two sample means. We can also use Fig. 7 for the operating characteristic of the test of significance of differences between means of two independent groups when the true standard deviations are equal. The test function is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{12}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}},$$

$$s_{12} = \sqrt{\frac{\Sigma(x_1 - \bar{x}_1)^2 + \Sigma(x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}},$$

where t has $n_1 + n_2 - 2$ degrees of freedom. Here $\bar{x}_1$ and $\bar{x}_2$ are the sample means and the sums of squares are taken for each sample separately.

Our hypotheses are

$$H_0: \ \mu_1 = \mu_2,$$
$$H_1: \ | \mu_1 - \mu_2 | = \lambda\sigma_1, \qquad (\lambda > 0),$$

where σ_1 is the unknown value of the standard deviation common to populations 1 and 2; the ratio of the standard deviations is unity ($k = 1$).

As before, we can take the effective sample size as

$$n' = \frac{n_1 n_2}{n_1 + n_2},$$

but for entering the figure we use

$$\lambda = \lambda' \sqrt{\frac{n'}{n}},$$

where $n = n_1 + n_2 - 1$. Note that we no longer have an arbitrary choice for n.

Example of two-sided t-test, for two sample means. If an attitude battery has standard deviation σ_1 roughly equal to 10, and the true difference between two ethnic group means is 5, how likely are we to accept the null hypothesis of no difference with samples of 25 and 50?

$$\lambda' = \frac{5}{10} = 0.5, \qquad n' = \frac{25 \times 50}{25 + 50} = 16.7,$$

but we do not enter Fig. 7 with these; we still need

$$\lambda = \lambda' \sqrt{\frac{n'}{n}} = 0.5 \sqrt{\frac{16.7}{74}} = 0.238, n = 74,$$

and we read the probability of accepting the null hypothesis as about 0.45.

SHORT-CUT ALLOWANCES IN THE ANALYSIS OF VARIANCE

In classical analysis of variance, we use the significance test (*F*-test) to decide whether or not all group means come from the same population. At best this is a trivial kind of question. Once we decide the groups are not identical, we want to know whether some specific groups have scored higher than others, and about this the classical analysis is strangely silent. In this section we present an easy way of making such comparisons. For general background, see *Analysis of Variance,* in Chapter 7.

The obvious way to compare groups in a one-way analysis of variance design is to make *t*-tests between pairs of groups, and this technique has often been used in the psychological literature. The shortcoming is that we may be fooling ourselves about the confidence level being used. Suppose we have 10 groups, all from the same population, then on the average it turns out that the largest observed mean and the smallest observed mean fall 3.1 standard errors ($\sigma/\sqrt{n}$) apart. A test of the difference of these extreme means gives quite a good chance of rejecting the equality of population means. If we use Fig. 6, we can regard each observation on a mean as one observation, and the standard deviations of these means can be regarded as 1. The effective sample size for the two means is

$$n' = 1 \times \frac{1}{1+1} = 0.5, \quad \lambda' = 3.1, \quad k = 1.$$

Choose $n = 1$, then $\lambda = 3.1\sqrt{0.5/1} = 2.2$, and we read that the probability of accepting H_0 is about 0.4, which is a far cry from the desired value of 0.95. This difficulty has prevented some authors from using all possible *t*-tests, while others have gone ahead, no doubt rightly feeling that there was little else they could do.

In the past few years several authors (Duncan, 1951, 1952; Tukey, 1949, 1951; Moshman, 1952; Link and Wallace, unpublished; Scheffé, 1953) have given attention to various ramifications of this problem. In 1953, we are a long way from seeing this problem wrapped up, but in this section we give one easy approach that may meet with favor. Following Tukey's lead, Link and Wallace developed some methods of setting allowances on differences between group means. These allowances are similar to confidence limits.

ONE-WAY CLASSIFICATION

The method of Link and Wallace can be used on one-way analysis of variance, and two-way analysis with one observation per cell. It uses ranges as a measure of variation rather than sums of squares. (A range is the largest minus the smallest observation.)

It seems not to be well known among social scientists that the range is quite a good measure of variation as long as sample sizes are reasonably small. As sample sizes get large the statistical efficiency of the range does decrease relative to the sample standard deviation. Nevertheless, it would be instructive to compare the accuracy of estimates of standard deviations computed from ranges and from sample standard deviations by a representative group of scientists. The thought is that the mistakes in adding, squaring, subtracting, dividing, and square rooting required by the sample standard deviation may actually contribute more serious discrepancies in routine calculations than is generally realized. Estimation from ranges seems to require fewer operations. Or do we all check our calculations so carefully that this thought can be put aside as unworthy?

More complicated designs require other approaches which we shall not attempt to discuss here. We illustrate with an example from Edwards (1950, p. 298-299) on maze learning under different instructions. In Table II we present the data and computations useful in the analysis. Note that there are three groups of seven members each.

TABLE II

ERRORS ON FIFTH TRIAL OF STYLUS MAZE FOR THREE GROUPS OF SEVEN SUBJECTS, EACH GROUP HAVING DIFFERENT INSTRUCTIONS

	Instructions		
	(1) "Intelligence test"	(2) "Average 20 errors"	(3) "Few errors as possible"
	20	20	29
	22	17	22
	26	26	23
	20	23	21
	26	18	32
	26	21	34
	21	22	24
Sum	161	147	185
Range	6	9	13
			Sum of ranges 28

We enter Table III with number of groups (=3) and number in group (=7) and observe the critical factor is 1.17. Multiplying the sum of the ranges by 1.17, we get 32.8. The importance of the number 32.8 is:

(i) If the totals of any two groups differ by more than 32.8 we "recognize" the existence of a real difference between the pair of groups.

(ii) The true difference between any two group totals is "recognized" as lying within ± 32.8 of that observed.

(iii) If we "recognize" as described we shall make one or more false "recognitions" in approximately 5 *experiments* out of 100.

For our example we "recognize" that the "Average 20 errors" instruction is different from the "Few errors as possible" instruction, but we do not "recognize" any other real differences. We can still set 95% allowances on the differences as follows:

Comparison	Diff. of totals	5% Allowance	Limits on totals	Limits on means
Instr. (1) — Instr. (2)	14	±33	−19 to 47	−2.7 to 6.7
Instr. (3) − Instr. (1)	24	±33	− 9 to 57	−1.3 to 8.1
Instr. (3) − Instr. (2)	38	±33	5 to 71	+0.7 to 10.1

The last column is obtained from the next to last by dividing by the number in a group. These limits are like confidence limits, in that we "recognize" that the true differences lie between these limits. At least one of these "recognitions" will be wrong in five experiments per 100, but all "recognitions" will be correct in 95 experiments per 100.

TWO-WAY CLASSIFICATION

Five judges rated affective value of 10 different colors on a nine-point rating scale (Guilford, 1936, p. 281). The data are presented in Table IV along with computations to assist in the allowance analysis.

In Table IV row and column totals have been computed as usual. The numbers in parentheses are the differences between adjacent columns, the last of which is the difference between the *E* and *A* columns or, in general, between the last and first columns. The range of the differences is just the range of the numbers appearing in parentheses in the same column. We could just as well have taken the differences the other way, between adjacent rows instead of between adjacent columns, but if one does

TABLE III

CRITICAL FACTORS FOR ALLOWANCES FOR ONE-WAY (BALANCED) DIVISION INTO GROUPS

5% RISK

Number in group = Number per range	Number of groups = number of ranges																		
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	3.43	2.37	1.78	1.40	1.16	1.00	.87	.78	.70	.66	.63	.58	.50	.47	.44	.42	.40	.38	.36
3	1.91	1.44	1.13	.94	.80	.70	.62	.56	.51	.47	.43	.40	.38	.36	.33	.32	.30	.29	.27
4	1.63	1.25	1.01	.84	.72	.63	.57	.51	.47	.43	.40	.37	.35	.33	.31	.29	.28	.27	.25
5	1.53	1.19	.96	.81	.70	.61	.55	.50	.45	.42	.39	.36	.34	.32	.30	.29	.27	.26	.25
6	1.50	1.18	.95	.80	.69	.61	.55	.49	.45	.42	.39	.36	.34	.32	.30	.29	.27	.26	.25
7	1.49	1.17	.95	.80	.69	.61	.55	.50	.45	.42	.39	.36	.34	.32	.30	.29	.28	.26	.25
8	1.49	1.17	.96	.81	.70	.62	.55	.50	.46	.42	.39	.37	.35	.33	.31	.29	.28	.27	.25
9	1.50	1.18	.97	.82	.71	.62	.56	.51	.47	.43	.40	.37	.35	.33	.31	.30	.28	.27	.26
10	1.52	1.20	.98	.83	.72	.63	.57	.52	.47	.44	.41	.38	.35	.34	.32	.30	.29	.27	.26
11	1.54	1.21	.99	.84	.73	.64	.58	.52	.48	.44	.41	.38	.36	.34	.32	.31	.29	.28	.27
12	1.56	1.23	1.00	.85	.74	.65	.59	.53	.49	.45	.42	.39	.37	.35	.33	.31	.30	.28	.27
13	1.58	1.25	1.02	.86	.75	.66	.59	.54	.49	.46	.42	.40	.37	.35	.33	.32	.30	.29	.27
14	1.60	1.26	1.03	.87	.76	.67	.60	.55	.50	.46	.43	.40	.38	.36	.34	.32	.31	.29	.28
15	1.62	1.28	1.05	.89	.77	.68	.61	.56	.51	.47	.44	.41	.38	.36	.34	.33	.31	.30	.28
16	1.64	1.30	1.06	.90	.78	.69	.62	.56	.52	.48	.44	.41	.39	.37	.35	.33	.31	.30	.29
17	1.66	1.31	1.08	.91	.79	.70	.63	.57	.52	.48	.45	.42	.39	.37	.35	.33	.32	.30	.29
18	1.68	1.33	1.09	.92	.80	.71	.64	.58	.53	.49	.46	.43	.40	.38	.36	.34	.32	.31	.30
19	1.70	1.34	1.10	.93	.81	.72	.65	.59	.54	.50	.46	.43	.40	.38	.36	.34	.33	.31	.30
20	1.72	1.36	1.11	.95	.82	.73	.65	.59	.54	.50	.47	.44	.41	.39	.37	.35	.33	.32	.30

Entries are to be multiplied by the sum of the ranges within groups to obtain allowances for group totals.

TABLE IV

RATINGS OF 10 COLORS BY 5 JUDGES

Color	*Judge* *A*	*B*	*C*	*D*	*E*	*Totals*
1	7 (3)	4 (−2)	6 (4)	2 (−3)	5 (−2)	24
2	8 (4)	4 (−3)	7 (5)	2 (−5)	7 (−1)	28
3	8 (4)	4 (−4)	8 (2)	6 (1)	5 (−3)	31
4	7 (2)	5 (−2)	7 (4)	3 (−1)	4 (−3)	26
5	5 (3)	2 (−5)	7 (4)	3 (−1)	4 (−1)	21
6	3 (2)	1 (−5)	6 (3)	3 (1)	2 (−1)	15
7	4 (1)	3 (−3)	6 (3)	3 (−3)	6 (2)	22
8	3 (0)	3 (−3)	6 (1)	5 (1)	4 (1)	21
9	6 (1)	5 (−3)	8 (4)	4 (−3)	7 (1)	30
10	4 (−2)	6 (−1)	7 (1)	6 (0)	6 (2)	29
Totals	55	37	68	37	50	247
Range of differences	6	4	4	6	5	
					Sum of ranges	25

both he should average the allowances obtained by the two approaches, rather than take the one with the smaller allowances.

We enter Table V with number of ranges = number of judges (=5) and number per range = number of colors (=10) and find the two entries 0.59, 0.49.

The upper entry corresponds to the column (judge) totals, while the lower entry corresponds to the row (color) totals. We multiply the two numbers by the sum of the ranges and get

0.59 × 25 = 14.8 = allowance for column totals.

0.49 × 25 = 12.3 = allowance for row totals.

Just as in the one-way classification, we can make comparisons of column totals and of row totals. We "recognize," for example, that Judge *C* differs from *B, D,* and *E,* but not from *A;* and we "recognize" that Color 6 differs from colors 2, 3, 9, and 10, but we "recognize" no other dif-

TABLE V

CRITICAL FACTORS FOR ALLOWANCES FOR TWO-WAY (BALANCED) DESIGN

Method—Successive Differences

5% RISK

Top entries lead to allowances for Variable *A* Totals
Bottom entries lead to allowances for Variable *B* Totals

Number of ranges = Number of Variable *A* totals

Number of differences per range = Number of Variable *B* totals	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	6.36 6.36	2.17 1.95	1.53 1.40	1.17 1.12	.96 .96	.80 .84	.68 .76	.62 .70	.54 .65	.49 .61	.45 .58	.41 .55	.38 .52	.36 .50	.34 .48	.31 .46	.30 .45	.28 .44	.27 .42
3	1.95 2.17	1.17 1.17	.90 .91	.73 .77	.62 .68	.54 .61	.48 .56	.42 .52	.38 .49	.35 .46	.32 .44	.30 .42	.28 .40	.26 .38	.25 .37	.23 .36	.22 .35	.21 .34	.20 .33
4	1.45 1.56	.96 .94	.76 .76	.63 .66	.54 .58	.47 .53	.42 .49	.38 .45	.34 .43	.32 .40	.29 .39	.27 .37	.25 .35	.24 .34	.23 .33	.21 .32	.20 .31	.19 .30	.18 .29
5	1.27 1.29	.90 .84	.72 .69	.60 .60	.51 .53	.45 .49	.40 .45	.36 .42	.33 .40	.30 .38	.28 .36	.26 .34	.25 .33	.23 .32	.22 .31	.21 .30	.20 .29	.19 .28	.18 .27
6	1.18 1.14	.87 .78	.70 .64	.58 .56	.50 .50	.44 .46	.39 .43	.36 .40	.33 .38	.30 .36	.28 .34	.26 .33	.24 .31	.23 .30	.22 .29	.21 .28	.20 .27	.19 .27	.18 .26
7	1.16 1.07	.85 .74	.69 .61	.58 .54	.50 .48	.44 .44	.39 .41	.36 .38	.33 .36	.30 .34	.28 .33	.26 .31	.24 .30	.23 .29	.22 .28	.21 .27	.20 .26	.19 .26	.18 .25
8	1.15 1.00	.85 .71	.69 .59	.58 .52	.50 .47	.44 .43	.40 .40	.36 .37	.33 .35	.30 .33	.28 .32	.26 .30	.25 .29	.23 .28	.22 .27	.21 .26	.20 .26	.19 .25	.18 .24
9	1.14 .96	.85 .68	.69 .57	.58 .50	.50 .45	.45 .42	.40 .39	.36 .36	.33 .34	.31 .33	.28 .31	.27 .30	.25 .29	.24 .28	.22 .27	.21 .26	.20 .25	.19 .24	.18 .24
10	1.15 .93	.86 .66	.70 .56	.59 .49	.51 .44	.45 .41	.40 .38	.37 .36	.34 .34	.31 .32	.29 .31	.27 .29	.25 .28	.24 .27	.23 .26	.21 .25	.20 .25	.19 .24	.19 .23

continued

ferences in color. We can also set limits on differences, as in the one-way analysis. All statements are made with a 5% error rate *per group of comparisons,* i.e., *per experiment.* Among all row (column) comparisons, we will be wrong at least once in five experiments per 100.

Although we cannot devote more space to this problem here, we remind the reader that much work is being done, and that further results should be made available to the psychological literature. Tables exist for 1% error rate, and methods exist for using classical analysis of variance computations. But the speed and simplicity of the present method go far to recommend it.

THE MATCHING PROBLEM

The matching problem occurs frequently in psychological investigations, but its existence is not widely recognized among social psychologists because it is not treated in many elementary texts and because it arises in subtle and unexpected places. For example, an experimenter presents pictures of faces and requests the subject to classify the pictures into their correct religious categories—Catholic, Jew, or Protestant. When the data are in, it is desired to test whether the number of correct responses hovers around that expected from guessing or

TABLE V (continued)

CRITICAL FACTORS FOR ALLOWANCES FOR TWO-WAY (BALANCED) DESIGN

Method—Successive Differences

5% RISK

Top entries lead to allowances for Variable *A* Totals
Bottom entries lead to allowances for Variable *B* Totals

Number of ranges = Number of Variable *A* totals

Number of differences per range = Number of Variable *B* totals	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
11	1.15	.87	.71	.60	.52	.46	.41	.37	.34	.31	.29	.27	.26	.24	.23	.22	.21	.20	.19
	.91	.65	.55	.48	.43	.40	.37	.35	.33	.31	.30	.29	.28	.27	.26	.25	.24	.24	.23
12	1.16	.88	.71	.60	.52	.46	.41	.38	.35	.32	.30	.28	.26	.25	.23	.22	.21	.20	.19
	.88	.63	.54	.47	.43	.39	.37	.35	.33	.31	.30	.28	.27	.26	.25	.25	.24	.23	.23
13	1.17	.89	.72	.61	.53	.47	.42	.38	.35	.32	.30	.28	.26	.25	.24	.22	.21	.20	.19
	.86	.62	.53	.47	.42	.39	.36	.34	.32	.31	.29	.28	.27	.26	.25	.24	.24	.23	.22
14	1.18	.90	.73	.62	.54	.48	.43	.39	.36	.33	.31	.29	.27	.25	.24	.23	.22	.21	.20
	.85	.62	.52	.46	.42	.39	.36	.34	.32	.30	.29	.28	.27	.26	.25	.24	.23	.23	.22
15	1.19	.91	.74	.63	.55	.48	.43	.39	.36	.33	.31	.29	.27	.26	.24	.23	.22	.21	.20
	.84	.61	.52	.46	.41	.38	.36	.33	.32	.30	.29	.28	.27	.26	.25	.24	.23	.23	.22
16	1.21	.92	.75	.64	.55	.49	.44	.40	.37	.34	.31	.29	.28	.26	.25	.23	.22	.21	.20
	.82	.60	.51	.45	.41	.38	.35	.33	.31	.30	.29	.27	.26	.25	.25	.24	.23	.23	.22
17	1.22	.93	.76	.65	.56	.50	.44	.40	.37	.34	.32	.30	.28	.26	.25	.24	.23	.22	.21
	.81	.60	.51	.45	.41	.38	.35	.33	.31	.30	.28	.27	.26	.25	.24	.24	.23	.22	.22
18	1.23	.94	.77	.65	.57	.50	.45	.41	.38	.35	.32	.30	.28	.27	.25	.24	.23	.22	.21
	.81	.59	.50	.45	.40	.37	.35	.33	.31	.29	.28	.27	.26	.25	.24	.24	.23	.22	.22
19	1.24	.95	.78	.66	.57	.51	.46	.42	.38	.35	.33	.31	.29	.27	.26	.24	.23	.22	.21
	.80	.59	.50	.44	.40	.37	.35	.33	.31	.29	.28	.27	.26	.25	.24	.23	.23	.22	.22
20	1.26	.97	.79	.67	.58	.51	.46	.42	.39	.36	.33	.31	.29	.27	.26	.25	.23	.22	.21
	.79	.58	.50	.44	.40	.37	.34	.32	.31	.29	.28	.27	.26	.25	.24	.23	.23	.22	.21

Entries are to be multiplied by the sum of ranges of differences between adjacent measurements to obtain allowances for Variable *A* totals (top entries) and Variable *B* totals (bottom entries).

whether the subject does get cues about religious affiliation from physical characteristics shown in pictures. Further examples are the following: (i) Two judges try to do the same content analysis on a set of stories, classifying each story by its main theme. Before we go further with their training, we stop to see whether there is any agreement at all between them. The number of stories they have called identically is the number of their matches. (ii) Two experimentally independent criteria are used to classify subjects (a group judgment on the basis of protocols and a projective test not given to the group), and it is desired to measure agreement using the number of identical classifications. (iii) Perhaps the most classical use of the matching problem is in extrasensory perception experiments, where the subject tries to name the face of a card not seen by him. Again the number of correct calls is a measure of the agreement. In all such procedures the preliminary question is whether the agreement is better than that expected from random guessing. If it is better, further questions of serious interest usually enter the discussion, such as exploration of factors that tend to increase the number of matches. We supply some results to assist in making the base-line computations (Battin, 1942; Wilks, 1943).

To facilitate discussion of this problem we shall call the stimulus materials the target deck of cards. The categories will be called suits. The number of separate stimuli in a particular category will be called the number of cards in the suit. The series of calls or sortings into categories by the subject will be called the matching deck. It too will have suits (usually the same as the target deck) and the numbers of cards in its suits will depend on the frequency with which each category is chosen. This implies that the subject is informed in advance of the number and nature of the suits.

The distribution of the number of correct calls in such experiments under a chance hypothesis can be obtained by counting as shown in the following example.

Example: Six handwriting specimens are available, three from convicted criminals whose crimes are against persons, and three from convicted criminals whose crimes are against property. A subject attempts to classify the handwriting into the two categories—three specimens to each. What is the chance distribution of the number of correct classifications? The first step is to lay out the target deck, and beneath it to write out all possible arrangements of the matching deck. We shall let *P* stand for property, *I* for persons (individuals).

Target	*P P P I I I*	*No. of correct matches*
Arrangements of matching deck	1. P P P I I I	6
	2. P P I P I I	4
	3. P P I I P I	4
	4. P P I I I P	4
	5. P I P P I I	4
	6. P I P I P I	4
	7. P I P I I P	4
	8. P I I P P I	2
	9. P I I P I P	2
	10. P I I I P P	2
	11. I P P P I I	4
	12. I P P I P I	4
	13. I P P I I P	4
	14. I P I P P I	2
	15. I P I P I P	2
	16. I P I I P P	2
	17. I I P P P I	2
	18. I I P P I P	2
	19. I I P I P P	2
	20. I I I P P P	0

If we take all matching arrangements as equally likely, then there is only one chance in 20 of getting all six correct. The mean and variance of the chance distribution of the number of correct matches can be obtained from the frequency distribution given below.

No. of matches: x	Frequency: f	fx	fx^2
6	1	6	36
4	9	36	144
2	9	18	36
0	1	0	0
Totals	20	60	216

$$\text{mean} = \bar{m} = \frac{60}{20} = 3,$$

$$\sigma = \sqrt{\frac{216}{20} - 3^2} = \sqrt{1.8} = 1.34.$$

Thus if 10 subjects average 4.0 correct matches we can test the chance hypothesis with the normal deviate

$$z = \frac{4-3}{\frac{1.34}{\sqrt{10}}} = \frac{\sqrt{10}}{1.34} = \frac{3.16}{1.34} = 2.36.$$

Consulting normal tables, we find this beyond the 5% level whether we are using a one-sided or a two-sided test. Actually, we might use a correction term here because subjects can get only an *even* number of correct matches. To do this properly we set up z based on the total number of correct matches:

$$z = \frac{|\,40-30\,| - 1}{\sqrt{18}} = \frac{9}{4.24} = 2.12,$$

which is still significant at the 5% level, whether one- or two-sided. The correction 1 is used only when there are exactly two suits. Otherwise the correction is ½.

We use the correction for continuity liberally in this chapter, so a word of explanation may be in order. Discrete distributions have their probability concentrated at points, and in most problems we discuss these points are integers: 0, 1, 2, . . . When we try to approximate such a discrete distribution by a continuous distribution such as the normal, the discrete density at a point such as 3 is usually thought of as being approximated by the area under the normal curve between 2.5 and 3.5. So if we want the probability of 3 or more heads in 4 tosses of a coin we would sum the probabilities of 3 and 4 heads from the discrete distribution to get $5/16 = 0.3125$. The expected number is $4 \times ½ = 2$ and the standard deviation is $\sqrt{4 \times ½ \times ½} = 1$. So the normal approximation is given by the area to the right of $(3 - ½) - 2 = ½$. This approximation is 0.3085. Had we not used the correction, the approximation would have been 0.1587, not nearly as close. In the present example the normal approximation is being made to a distribution that takes only even values, so the probability at 40 matches is represented by the area of the normal from 39 to 41, so in this case the correction term is 1. It might be remarked that the correction is being made on the normal deviate to adjust for the discontinuity of the discrete distribution. Such a correction could just as well have been called a correction for discontinuity.

We will not usually have the exact distribution of the number of matches at our disposal. (It is easy to see that computations of the exact distribution will become extremely laborious as the number of cards and suits gets large.) Consequently, we shall have to rely on approximate normality, and compute means and standard deviations for approximate t-tests.

Identical target and matching decks, same number of cards in each suit. Suppose there are s suits, with c cards per suit in both the target and the matching deck. The total number of cards per deck is $n = sc$. Then the average and variance of the number of matches under chance arrangement of the matching deck are

$$\bar{m} = c = np, \quad where\ p = \frac{1}{s}, \quad q = 1 - p,$$

$$\sigma^2 = \frac{n^2\,(s-1)}{s^2\,(n-1)} = npq\left(\frac{n}{n-1}\right)$$

Here p is the probability that the first card is put in the correct suit. Because there are s suits of equal size, the chance is $1/s$. This elegant way of writing the mean and variance in terms of n, p, and q was pointed out by C. Horace Hamilton. It shows us that, properly interpreted, the mean and variance are approximately those of a binomial. The mean is exactly in binomial form, while the binomial variance npq is multiplied by a factor $n/(n-1)$ which is very close to unity when n is large.

If we apply these results to the handwriting example, we have

$$s = 2,\ c = 3,$$

$$\bar{m} = c = 3,$$

$$\sigma^2 = npq\left(\frac{n}{n-1}\right) = 6 \times (½)\,(½)\,(6/5)$$
$$= 9/5 = 1.8,$$

$$\sigma = \sqrt{1.8} = 1.34.$$

These results agree with our more detailed calculations.

Identical target and matching decks, differing numbers of cards per suit. By design we may not have the same number of stimulus materials in each category. Things become more complicated. Suppose there are s suits with c_i cards in the ith suit.

$$n = \sum_{i=1}^{s} c_i\ ,$$

$$\bar{m} = \frac{\sum_{i=1}^{s} c_i^2}{n},$$

$$\sigma^2 = \frac{1}{n-1}\left[\bar{m}\,(\bar{m} + n) - \frac{2}{n}\sum_{i=1}^{s} c_i^3\right].$$

Example: The handwriting expert has three classes of criminals to classify. The numbers actually in the three classes are 5, 3, and 2. He is told the number in each class and asked to put these numbers into the classes. Then

$$n = 5 + 3 + 2 = 10,$$

$$\bar{m} = \frac{25 + 9 + 4}{10} = \frac{38}{10} = 3.8,$$

$$\sigma^2 = \tfrac{1}{9}\,[3.8\,(3.8 + 10) - (2/10)\,(125 + 27 + 8]$$

$$= \frac{20.44}{9} = 2.27.$$

If the expert gets 6 correct our normal deviate is

$$z = \frac{6 - 3.8 - 0.5}{\sqrt{2.27}} = \frac{1.7}{1.51} = 1.13.$$

The 0.5 is the continuity correction. A deviation of 1.13 is not significant at the 0.05 level (one-sided).

The actual table of data might have been shown in a 3 × 3 table, as we illustrate.

3 × 3 TABLE FOR HANDWRITING EXAMPLE

Target deck
Categories

		1	2	3	Totals
Calls	1	3	2	0	5
	2	2	1	0	3
	3	0	0	2	2
Totals		5	3	2	10

The number of matches is the sum of the diagonal cells: $3 + 1 + 2 = 6$.

Nonidentical target and matching decks, target deck having suits of same size. Sometimes, unfortunately, the number of times a subject calls each suit cannot be controlled. In such situations the computations are still fairly simple if each suit of the target deck has the same number of cards.

Let s be the number of suits, and c the number of cards per suit of the target deck. The matching deck (the calls) has $c_1, c_2, \ldots, c_s$ cards in the suits, $\Sigma c_i = n$. Then

$$\bar{m} = c,$$

$$\sigma^2 = \frac{c[n^2 - \Sigma c_i^2]}{n\,(n - 1)}.$$

Example: We have three categories for handwriting, with 5 specimens of each. The subject puts 7 in category 1, 5 in category 2, and 3 in category 3. Then

$$c_1 = 7,\ c_2 = 5,\ c_3 = 3,\ n = 15.$$

$$\bar{m} = 5,$$

$$\sigma^2 = \frac{5\,(225 - [49 + 25 + 9])}{15\,(14)}$$

$$= \frac{142}{3\,(14)} = 3.38.$$

Had the subject put the same number in each category, the mean would still be 5, but the variance would be

$$\frac{225\,(\tfrac{1}{3})\,(\tfrac{2}{3})}{14} = \frac{50}{14} = 3.57$$

instead of 3.38. In general, the use of unequal numbers of calls per category when the target deck has equal numbers will reduce the variance. This means that in testing significance for the mean scores of several subjects the use of the equal-category variance will give a conservative test in cases where the experimenter is unwilling to compute the variance for each subject separately.

Here the subject has probably called freely. We have tested his accuracy, given the particular number of calls in each category he has made. The question is sometimes raised whether the subject should not get credit for accurately reproducing the correct number of calls in each category. It is difficult to conceive of a probability model that would accurately reproduce the physical situation. We cannot tell what suspicions the subject has about our category sizes, or from what kind of distribution he chooses his calls. It seems difficult therefore to do much about this aspect of the problem.

Different decks. In special cases the matching problem will occur with target and matching decks of quite different compositions. Let the number of cards in each suit of the target deck be $t_1, t_2, \ldots, t_s$ and, as before, the number of cards in the suits of the matching deck be $c_1, c_2, \ldots, c_s$. Then $n = \Sigma c_i = \Sigma t_i$. The mean number of correct matches is

$$\bar{m} = \frac{\Sigma t_i c_i}{n},$$

$$\sigma^2 = \frac{1}{n^2\,(n - 1)}\,[\,(\Sigma t_i c_i)^2 - n\,\Sigma t_i c_i\,(t_i + c_i) + n^2\,\Sigma t_i c_i].$$

Example: In an open-end pilot study of classification of emotions from photographs, a student has 20 photographs: 5 expressing "anger," 7 expressing "joy," 8 expressing "sorrow." The first subject gives 3 responses of "joy," 4 "sorrow," 6 "fear," 2 "tranquility," 5 "I don't know." The total number of categories is obtained by counting the number of differently named categories that are used in one deck, the other, or both. There are 6. The set-up shown below for "Computations for nonidentical decks" will facilitate the calculations.

Suppose the subject gets 5 correct matches, is he doing better than chance at approximately the 5% level, considering the categories he used? We should probably use the continuity correction of ½ for his score, because he can get only an integral number of matches. Then

$$z = \frac{(5-\frac{1}{2}) - 2.65}{\sqrt{1.596}} = \frac{1.85}{1.26} = 1.47.$$

Because the normal table gives 1.65 for 5% one-tailed test we would not reject the hypothesis of chance sorting on the basis of these data. This subject could have gotten at most seven correct matches. Although the sensitivity (or power) of the matching test has not been worked out, it is clear that a situation that requires a subject to get 6 out of 7 agreements for significance does not give the subject much chance to display his ability if he has only a little. This is why open-end responses in matching situations are usually not desirable; they cut down the possibility of a significant result for a single subject by decreasing the number of possible matches. Furthermore, each subject (especially when the target suits have unequal sizes) requires a special computation for his mean and variance. Suppose we have several subjects in the experiment just described with the following data:

Subject	No. of correct matches	$\overline{m}$	σ^2
1	5	2.65	1.60
2	4	2.00	1.45
3	6	2.34	1.36
4	7	6.51	1.00
Totals	22	13.50	5.41

The significance test would be

$$z = \frac{(22-\frac{1}{2}) - 13.50}{\sqrt{5.41}} = \frac{8.00}{2.33} = 3.43,$$

significant well beyond the 5% level. *Note:* (1) the ½ correction is made on the total number of hits, (2) the hypothetical mean is the sum of the means, (3) the hypothetical variance is the sum of the variances.

From the discussion the reader will note that the application of the test of significance for matching is the same whether the experimenter informs his subjects of the number of suits and number within a suit, or whether he tells his subjects nothing about the target deck. It is usually preferable for the experimenter to give his subjects complete information about the composition of the target deck because, as can be seen from the examples, this gives greater chance for subjects to display their ability at matching if they have any. A possible exception to this general rule is the case where an experimenter is conducting a pilot study, not only to get some idea about matching ability, but also some information about what "natural" categories might exist for his particular matching problem.

COMPUTATIONS FOR NONIDENTICAL DECKS

	Anger	Joy	Sorrow	Fear	Tranquility	DK	Totals
Photographs t_i	5	7	8	0	0	0	20
Responses c_i	0	3	4	6	2	5	20
$t_i + c_i$	5	10	12	6	2	5	
$t_i c_i$	0	21	32	0	0	0	53
$t_i c_i (t_i + c_i)$	0	210	384	0	0	0	594

$$\overline{m} = \frac{53}{20} = 2.65,$$

$$\sigma^2 = \frac{1}{20^2(19)}[(53)^2 - 20(594) + 20^2(53)] = \frac{12,129}{7,600} = 1.596.$$

NONPARAMETRIC METHODS

The following discussion of nonparametric methods is drawn almost entirely from Moses (1952), with the exception of the discussion of Whitney's extension of the U-test, the analysis of variance for ranks, the tables, and some opinions of our own. Moses states (pp. 123-124) the advantages and disadvantages of nonparametric methods as follows:

> *Advantages of nonparametric methods:*
>
> 1. Whatever may be the form of the distribution from which the sample has been drawn, a nonparametric test of a specified significance level actually *has* that significance level (provided that the sample has been drawn at random; in certain cases, as will be noted, it is also necessary to assume that the distribution is continuous).
> 2. If samples are very small, e.g., six, there is in effect no alternative to a nonparametric test (unless the parent distribution really is known).
> 3. If the sample consists of observations from several *different* populations there may be a suitable nonparametric treatment.
> 4. The methods are usually easier to apply than the classical techniques.
> 5. If the data are inherently of the nature of ranks, not measurements, they can be treated directly by nonparametric methods without precariously assuming some special form for the underlying distribution.
> 6. In certain cases data can only be taken as "better" or "worse," that is, an observation can only be characterized as a plus or minus. Obviously, the classical tests are not directly applicable to such data.
>
> *Disadvantages of nonparametric methods:*
>
> 1. If nonparametric tests rather than normal-theory tests are applied to normal data then they are wasteful of data. The degree of wastefulness is measured by the "efficiency" of the nonparametric test. If, for example, a test has 80 percent efficiency, this means that *where the data are from a normal distribution,* the appropriate classical test would be just as effective with a sample of 20 percent smaller size. The efficiency thus expresses the relative merits of the nonparametric test and the classical test under the conditions *where the normal test is correct,* but does not tell us how the tests will compare on non-normal data.
> 2. The nonparametric tests and tables of significance values are widely scattered in the periodical literature.
> 3. For large samples some of the nonparametric methods require a great amount of labor, unless approximations are employed.

It might help to translate disadvantage 1 into the same terms we used to discuss the power of some of the common statistical tests. If we think of how likely we are to detect a true difference when it exists, to say the nonparametric method is less powerful than the test based on normal theory is to say that we will be less likely to detect a true difference of a given size. On the other hand, this is a two-edged sword, for if observations are cheap and computation using normal theory is expensive, it may be efficient in the ordinary sense of the word rather than the statistical sense to use nonparametric methods anyway.

The sign test. This test gets its name from the fact that we replace attribute or measurement data with plus and minus signs, and use the binomial distribution with $p = q = \frac{1}{2}$ as the null hypothesis. Suppose we have Treatment A and Treatment B; we are interested in seeing whether these two treatments differ. Values observed under Treatment B are subtracted from values observed under A, and then, depending on whether this difference is positive or negative, we write down a plus sign or a minus sign. If there are ties, throw away all tied pairs and work with the reduced sample size. [Although this is not the usual recommendation, it seems appropriate from an examination of Tocher's handling of Fisher's Flower Problem (1950).] The assumptions underlying this test are (1) the variable under consideration has a continuous distribution, (2) that both members of a pair are treated similarly except for getting Treatment A or Treatment B. Notice that assumption (2) does not imply similar treatment of *various* pairs. The hypothesis being tested is that the population medians are equal. Under the null hypothesis, we expect roughly equal numbers of plus and minus signs, and the probability level of a particular result can be assessed using the binomial distribution with n = number of pairs, and $p = q = \frac{1}{2}$. There are tables of the binomial distribution for n up to 49 (Tables, 1950) and, of course, even without the tables we can use the normal approximation to the binomial (with the correction):

$$z = \frac{|x - np| - \frac{1}{2}}{\sqrt{npq}} = \frac{|x - \frac{1}{2}n| - \frac{1}{2}}{\frac{1}{2}\sqrt{n}},$$

where x is the observed number of plus (minus) signs, n is the number of pairs, and $\frac{1}{2}$ is the continuity correction. This test can easily be extended to cases where the question is slightly reformulated: Is A better than B by 5 points, or is A better than B by 10%? In order to do this, we merely add 5 points or 10% to the values of the B's before proceeding with taking differences. It is easy to see that we can manipulate the formula to suit a large variety of hypotheses.

To facilitate the use of the sign test we include a table of critical values of x for the significance levels 1%, 5%, 10%, 25%. The tabulation is planned for use in a two-sided test. If

$$r = \text{smaller of } x \text{ and } n - x,$$

then we merely look in Table VI to see whether the observed r is less than or equal to the tabu-

TABLE VI

TABLE OF CRITICAL VALUES OF r FOR THE SIGN TEST

	Percent level of significance			
n	1	5	10	25
1	—	—	—	—
2	—	—	—	—
3	—	—	—	0
4	—	—	—	0
5	—	—	0	0
6	—	0	0	1
7	—	0	0	1
8	0	0	1	1
9	0	1	1	2
10	0	1	1	2
11	0	1	2	3
12	1	2	2	3
13	1	2	3	3
14	1	2	3	4
15	2	3	3	4
16	2	3	4	5
17	2	4	4	5
18	3	4	5	6
19	3	4	5	6
20	3	5	5	6
21	4	5	6	7
22	4	5	6	7
23	4	6	7	8
24	5	6	7	8
25	5	7	7	9
26	6	7	8	9
27	6	7	8	10
28	6	8	9	10
29	7	8	9	10
30	7	9	10	11
31	7	9	10	11
32	8	9	10	12
33	8	10	11	12
34	9	10	11	13
35	9	11	12	13
36	9	11	12	14
37	10	12	13	14
38	10	12	13	14
39	11	12	13	15
40	11	13	14	15
41	11	13	14	16
42	12	14	15	16
43	12	14	15	17
44	13	15	16	17
45	13	15	16	18
46	13	15	16	18
47	14	16	17	19
48	14	16	17	19
49	15	17	18	19
50	15	17	18	20
51	15	18	19	20
52	16	18	19	21
53	16	18	20	21
54	17	19	20	22
55	17	19	20	22
56	17	20	21	23
57	18	20	21	23
58	18	21	22	24
59	19	21	22	24
60	19	21	23	25
61	20	22	23	25
62	20	22	24	25
63	20	23	24	26
64	21	23	24	26
65	21	24	25	27
66	22	24	25	27
67	22	25	26	28
68	22	25	26	28
69	23	25	27	29
70	23	26	27	29
71	24	26	28	30
72	24	27	28	30
73	25	27	28	31
74	25	28	29	31
75	25	28	29	32
76	26	28	30	32
77	26	29	30	32
78	27	29	31	33
79	27	30	31	33
80	28	30	32	34
81	28	31	32	34
82	28	31	33	35
83	29	32	33	35
84	29	32	33	36
85	30	32	34	36
86	30	33	34	37
87	31	33	35	37
88	31	34	35	38
89	31	34	36	38
90	32	35	36	39
91	32	35	37	39
92	33	36	37	39
93	33	36	38	40
94	34	37	38	40
95	34	37	38	41
96	34	37	39	41
97	35	38	39	42
98	35	38	40	42
99	36	39	40	43
100	36	39	41	43

For $n > 100$, approximate values of r may be found by taking the nearest integer less than $\frac{1}{2}n - k\sqrt{n}$, where $k = 1.3, 1, .82, .58$ for the 1, 5, 10, 25 percent values respectively. A closer approximation to the values of r is obtained from $\frac{1}{2}(n - 1) - k\sqrt{n + 1}$ and the more exact values of k, 1.2879, .9800, .8224, .5752.

lated one. For example, with $n = 23$ observations we observe 5 +'s and 18 —'s. Then $r = 5$. We find this number significant at the 5%, 10% and 25% levels, but not at the 1% level. Had the direction of the difference been predicted, we would have used a one-sided test and halved the percentages at the top of the columns. In one-sided tests we must be careful that the observed result was in the direction of the alternative hypothesis.

A table related to the power of the sign test is also helpful. Table VII shows the sample size needed in two-sided tests at various levels of significance to be very sure (95%) of rejecting the null hypothesis $p = \frac{1}{2}$, when it is indeed false. For example, if the significance level is 5% and the actual p is 0.40, we need a sample of about 327, and we reject if x or $n - x$ is less than or equal to 145. Alternatively, suppose it is planned to use a sample of size 50, the true p would have to deviate from 0.50 by 0.25 before we would be very sure of getting significance at the 5% level. Like the earlier curves of the power of the common tests of significance, Table VII deserves more serious attention than it has ordinarily been given. It suggests that if deviations of p from 0.50 of the order of 0.05 are being sought, we need samples of the order of 1000; for deviations of 0.10, we need samples of the order of 300; and it is not until huge deviations such as 0.35 occur that ordinarily used samples such as $n = 25$ become reasonably adequate.

TABLE VII

MINIMUM VALUES OF SAMPLE SIZE n AND CRITICAL VALUES r, NECESSARY TO FIND SIGNIFICANT DIFFERENCES 95% OF THE TIME FOR VARIOUS GIVEN PROPORTIONS

	n = *Sample size*				r = *Critical values*			
p_0	$\alpha =$ 1%	5%	10%	25%	$\alpha =$ 1%	5%	10%	25%
.45 (.55)	*1,777*	*1,297*	*1,080*	*780*	*833*	*612*	*512*	*373*
.40 (.60)	*442*	327	*267*	*193*	*193*	145	*119*	*87*
.35 (.65)	*193*	143	*118*	86	*78*	59	*49*	37
.30 (.70)	*106*	79	67	47	*39*	30	26	19
.25 (.75)	66	49	42	32	22	17	15	12
.20 (.80)	44	35	28	21	13	11	9	7
.15 (.85)	32	23	18	14	8	6	5	4
.10 (.90)	24	17	13	11	5	4	3	3
.05 (.95)	15	12	11	6	2	2	2	1

The italicized values are approximate. The maximum error is about 5 for the value of n, and 2 for the value of r. The values of n and r for 5% were taken from MacStewart (1941), who gives a table of values of n and r for a range of confidence coefficients (the above table uses only 95%) and a single value $\alpha = 5\%$.

The median test. Sometimes we wish to phrase the question we ask as follows: Do two samples, n X's and m Y's, have the same median? There do not have to be equal numbers of observations in the two sets. First, compute the median for the pooled samples. If the null hypothesis is true and the two sets of observations do have the same median, we should find roughly half the X's and half the Y's above the combined median.

To carry out the test, record all the observations falling above the common median as +'s, and all falling below as —'s. When the total number of observations, $n + m$, is odd, one observation will fall at the median; this observation should be thrown away. Then record the number of +'s and —'s in each sample in a 2×2 table:

	+'s	—'s	
X's	r	$n - r$	n
Y's	s	$m - s$	m
	$r + s$	$n + m - r - s$	$n + m$

Make the usual chi-square test for independence with one degree of freedom, using Yates' correction. The question we are asking when we make this test is: Is the attribute of being an X associated with the attribute of being above the common median?

Moses says (p. 126): "The assumptions underlying this test are that the X's and Y's are random samples from their respective populations, and that the population distributions are of the same form, differing only by a translation up or down the scale. Although the test is derived using the second assumption, Mood states (1950, p. 394) that the test 'is sensitive primarily to differences in location and very little to differences in shape.'" The cutting point need not be the median of the combined sample but should be chosen before looking at the data. As a general rule, use of the median should give a more powerful test, so it might as well be used consistently.

Wilcoxon's matched pairs test using signed ranks. This test uses the same information the sign test uses, and in addition takes some account of the *size* of the observed differences. This means it is a more *powerful* test than the sign test (smaller probability of a type II error), because it uses more information.

The first step is to take differences by pairs, and then rank these differences in order of absolute size (disregarding sign). The smallest numerical difference is ranked as 1. Sum separately the ranks of the positive and of the negative differences; use the smaller sum. Suppose we have 7 pairs, with the values shown in the example.

WILCOXON SIGNED RANK EXAMPLE

Pair	Treatment A	Treatment B	Difference $A-B$	Rank of absolute difference	Ranks of same sign with lower total
1	65	51	14	6	
2	60	44	16	7	
3	71	64	7	4	
4	52	57	–5	2	2
5	62	49	13	5	
6	41	38	3	1	
7	58	52	6	3	——
			Sum of ranks with the lower total		2

If the sum of the ranks for either the negative differences or the positive differences is too small, the null hypothesis is rejected. As mentioned above, we take the sum of the ranks of the observations with smaller rank total.

We give tables of critical values of the sum of the ranks, T, for n from 7 to 16. For values of n larger than 16 we suggest the approximate normal deviate

$$z = \frac{T - \overline{T} + \frac{1}{2}}{\sigma_T},$$

where

$$\overline{T} = \frac{n(n+1)}{4},$$

$$\sigma_T{}^2 = \frac{(2n+1)\,\overline{T}}{6},$$

and the $\frac{1}{2}$ is a continuity correction. Note that when we reject the null hypothesis T will be less than $\overline{T}$ and therefore z will be negative.

The data presented above, when tested using the sign test, show no difference between Treatments A and B. However, when signed ranks are used, the conclusion is that Treatment A is significantly different from Treatment B at the 5% level (see Table VIII). This happens because the smallness (—5) of the one negative difference adds its weight to the fact that only this one negative difference is observed. It is obvious from the way the test is carried out that one negative difference with rank 6 will count just as heavily against the hypothesis as three negative differences with ranks 1, 2, and 3. Readers interested in setting confidence limits using signed ranks should see Moses (1952).

TABLE VIII

WILCOXON SIGNED RANK TEST

Given matched pairs of observations. Rank differences in order of increasing absolute *magnitude.* Compute the sum of ranks for positive differences and for negative differences. Let n equal the number of pairs, T be the smaller sum of ranks, and P be the probability of obtaining a given T or less by chance.

n	T	P
7	0	0.016
7	2	0.047
8	0	0.0078
8	2	0.024
8	4	0.055
9	2	0.0092
9	3	0.019
9	6	0.054
10	3	0.0098
10	5	0.019
10	8	0.049
11	5	0.0093
11	7	0.018
11	11	0.053
12	7	0.0093
12	10	0.021
12	14	0.054
13	10	0.0105
13	13	0.021
13	17	0.050
14	13	0.0107
14	16	0.021
14	21	0.054
15	16	0.0103
15	19	0.019
15	25	0.054
16	19	0.0094
16	23	0.020
16	29	0.053

Mann-Whitney test. The parallel nonparametric test to use with ranks when two samples are not composed of matched pairs, and may be of unequal size, is the Mann-Whitney test. [Wilcoxon (1945) seems to have invented this test too, as Mann and Whitney note, but the latter authors made a rather extensive investigation of its properties. Because of its elegant simplicity it is understandable that it has been invented independently quite often.] The null hypothesis is that the two samples have been drawn from a common population. If there are two populations *A* and *B* the test is sensitive to differences in the population that make it more likely that an observation drawn from *A* will be larger than one from *B*. An easy example is that of two populations identical in shape except that their means differ.

We suppose there are *X*'s and *Y*'s, *m* of one kind and *n* of the other. The first step is to arrange the $(m + n)$ observations in order from least to greatest, and then to give the observations their ranks. For example, suppose we have samples of *X*'s and *Y*'s as follows:

X: 1, 8, 17, 22, 45
Ranks: 1, 5, 9, 10, 11

Y: 2, 6, 7, 9, 10, 12
Ranks: 2, 3, 4, 6, 7, 8

Sum of ranks of smaller sample
$T = 36$

Having ranked the observations, we sum the ranks in the sample of smaller size (if the sizes are equal it does not matter which is chosen for a two-sided test). We call this sum *T*, and in our example its value is 36. The size of the smaller sample is called *m*, and the size of the larger *n*. In our example $m = 5$, $n = 6$. To test significance we enter Table IX with these values of *m* and *n*. Because 36 does not fall on or outside the limits 18 and 42 given there, we will not reject at the 5% level the hypothesis that the populations are identical. Table IX was derived from Mann and Whitney's tables (1947) of the distribution of their *U*-statistic. The statistic *T* seems slightly easier to use because it involves less computation than *U*. For one-sided tests, where the directionality of the difference is suggested in advance of the data, the significance levels in Table IX can be halved.

With pairs of samples that fall outside Table IX, provided *m* is greater than four, we should obtain a close approximation to the correct significance level by using as a normal deviate

$$z = \frac{|T - \overline{T}| - \frac{1}{2}}{\sigma_T},$$

where

$$\overline{T} = \frac{m}{2}(m + n + 1),$$

$$\sigma_T^2 = \frac{mn(m + n + 1)}{12}.$$

It is clear that $\overline{T}$ has the form given because the average rank in the whole set is $(m + n +$

TABLE IX

MANN-WHITNEY TEST

Critical values of *T*, the sum of the ranks for the smaller sample. The smaller sample has size *m*, and the larger one size *n*. The hypothesis of equality of distributions is rejected if the observed *T* falls at either limit, or outside the range of the pair of numbers shown in each cell. The upper pair gives critical limits at the 5% level (2-sided), and the lower pair in parentheses gives the critical limits at the 10% level (2-sided).

$n \backslash m$	2	3	4	5	6	7	8
3	— —	— (6,15)					
4	— —	— (6,18)	10,26 (11,25)				
5	— (3,13)	6,21 (7,20)	11,29 (12,28)	17,38 (19,36)			
6	— (3,15)	7,23 (8,22)	12,32 (13,31)	18,42 (20,40)	26,52 (28,50)		
7	— (3,17)	7,26 (8,25)	13,35 (14,34)	20,45 (21,44)	27,57 (29,55)	36,69 (39,66)	
8	3,19 (4,18)	8,28 (9,27)	14,38 (15,37)	21,49 (23,47)	29,61 (31,59)	38,74 (41,71)	49,87 (51,85)

1) /2, and there are m things ranked. The variance of T is also obvious from the formula for the variance of sums of observations drawn from a finite population, namely, $m\sigma^2(N - m)/(N - 1)$, where N is the population size, m is the sample size, and σ^2 is the variance of the original finite population. Here $N = m + n$, m is our smaller sample size, σ^2 is the variance of the first $m + n$ integers, which is well known to be $[(m + n)^2 - 1]/12$. Making these substitutions gives $\sigma_T{}^2$ immediately.

Whitney's extension of the U-statistic. When there are three samples rather than two, we may wish a nonparametric test to decide whether all three came from the same population, or whether to prefer an alternative such as

(i) a particular population, specified in advance, tends to yield smaller observations than either of the others; or

(ii) a particular order for the populations, specified in advance, exists such that if one observation is drawn from each, and any particular number is chosen, the observation from the population A is more likely to be smaller than the chosen number than the observation from population B, and the observation from population B is more likely to be smaller than the chosen number than the observation from population C.

If the frequency distributions A, B, and C do not overlap and their means are 10, 50, 120, respectively, then they are clearly in the order A, B, C. When the frequency distributions do overlap it becomes necessary to introduce the words "stochastically smaller." Perhaps a numerical example will make this idea clearer. We shall work with test scores on four groups, A, B, C, D. The main point concerns the cumulative frequency distributions, and that is what we give below.

Cumulative frequency distributions of four groups of test scores. Entry is percent of group scoring at or below the number in the "test score" column.

	Groups			
Test score	*A*	*B*	*C*	*D*
10	100	100	100	100
9	100	100	100	95
8	100	100	100	80
7	100	100	90	75
6	100	90	80	60
5	100	75	50	55
4	100	40	30	30
3	75	25	15	25
2	50	20	10	20
1	20	15	5	18
0	10	5	0	8

In the table the scores from group A are stochastically smaller than those for B, C, or D because group A always has at least as many scores at or below a particular number as any of the other groups. Similarly, the scores from group B are stochastically smaller than those from group C. Note that it is the test scores that tend to be smaller, the entries in the table are larger. Group D scores are neither stochastically larger nor stochastically smaller than those for groups B and C. This happens because group D has too many high scores and too many low scores. Comparing D with B, for scores from 4 to 9, B has the more frequent small scores, but for scores 0 and 1, D has the more frequent. We return now to the alternatives given above.

If alternative (i) were true we would say that observations from the first population were stochastically smaller than those for either of the other two populations. If alternative (ii) were true we would say that observations from population A were stochastically smaller than those from the other two, and that observations from population B were stochastically smaller than those from C. An easy example occurs when we have three normal populations with equal standard deviations but different means. Then observations from the population with smallest mean are stochastically smaller than observations from either of the other populations.

Our interest in a test with alternatives like (i) and (ii) is in the spirit of our discussion of allowances earlier. We are not satisfied to discover whether the three populations differ, we want to know whether the results fall in an order specified in advance. When we have a theory about three groups, the theory will usually speak about the order of the outcomes. We want to have a statistical test to go with such specific predictions. It is some trouble to work out the test, but if one has this specific problem he may find it worth the effort. We have provided tables to facilitate the work.

Suppose we have the following setup:

Populations	A	B	C
Samples	x	y	z
Sample sizes	l	m	n

We deal first with alternative (i). Are the x's stochastically smaller than the y's and the z's? If the x's are smallest in the population, they should exhibit it in the sample. We define two quantities U and V. The first is a measure of how often y's are smaller than x's, and the second a measure of how often z's are smaller than x's. Clearly, if the x's are stochastically smallest, U and V should be small. Specifically:

> U is the number of times a y is smaller than an x. Each y is tried against each x.
>
> V is the number of times a z is smaller than an x. Each z is tried against each x.

Example: When the observations are arranged from least to greatest, we observe for 6 x's, 3 y's and 3 z's the order

$$y\,y\,x\,x\,x\,y\,x\,z\,x\,z\,z\,x.$$

The first and second y's are ahead of 6 x's each and the third y is ahead of 3 x's, making a total of $U = 15$. Similarly, the first z is ahead of two x's and the last two z's are ahead of one each, making a total $V = 4$.

When the null hypothesis is true and the sample sizes are not too small (it is not actually known what too small is; we would advise samples of at least 5), it turns out that U and V are approximately distributed according to a normal bivariate distribution with means

$$\overline{U} = lm/2 \qquad \overline{V} = ln/2$$

and variances

$$\sigma_U^2 = \frac{lm\,(l+m+1)}{12}$$

$$\sigma_V^2 = \frac{ln\,(l+n+1)}{12},$$

and correlation

$$\rho_{UV} \doteq \sqrt{\frac{mn}{(l+m+1)\,(l+n+1)}}.$$

The observed U and V can be transformed to standard normal deviates:

$$u = \frac{U-\overline{U}}{\sigma_U}, \qquad v = \frac{V-\overline{V}}{\sigma_V},$$

and we ask whether u and v are both too small.

To make significance tests we have to use Pearson's tables of the bivariate normal (Pearson, 1931). The tables give for selected pairs of values of standardized normal deviates h, k with various correlations, the probability that both h and k will be exceeded. This means that we should be able to set the size of the region for rejection of the null hypothesis by using the tables. If the sample sizes l, m, n are roughly equal, then to test whether x is the smallest we want simultaneously

$$u = \frac{U-\overline{U}}{\sigma_U} \leqq h, \qquad v = \frac{V-\overline{V}}{\sigma_V} \leqq k,$$

where h and k are chosen to give an appropriate significance level. Whitney suggests that h and k might be set equal, the critical value being a value of h such that the probability that both inequalities hold when the null hypothesis is true is about 5%.

Since we expect U and V to be small, the values of h and k will be negative. Table X gives the values of h for various correlation coefficients ρ at three significance levels: 1%, 5%, 10%, i.e.,

$$\text{Prob.}\,(u \leqq h,\ v \leqq h) = 0.01,$$

$$\text{Prob.}\,(u \leqq h,\ v \leqq h) = 0.05,$$

$$\text{Prob.}\,(u \leqq h,\ v \leqq h) = 0.10.$$

TABLE X

VALUES FOR h FOR TESTING NULL HYPOTHESIS AGAINST ALTERNATIVE HYPOTHESIS

$x < y, \qquad x < z$

	Significance level		
Correlation	0.10	0.05	0.01
0.00	–0.48	–0.76	–1.28
0.05	–0.51	–0.79	–1.33
0.10	–0.54	–0.83	–1.37
0.15	–0.57	–0.86	–1.41
0.20	–0.59	–0.89	–1.46
0.25	–0.62	–0.93	–1.50
0.30	–0.65	–0.96	–1.54
0.35	–0.68	–1.00	–1.59
0.40	–0.71	–1.03	–1.63
0.45	–0.74	–1.07	–1.67
0.50	–0.77	–1.10	–1.71
0.55	–0.80	–1.14	–1.76
0.60	–0.84	–1.18	–1.80
0.65	–0.87	–1.21	–1.85
0.70	–0.91	–1.25	–1.89
0.75	–0.95	–1.29	–1.95
0.80	–0.99	–1.34	–1.99
0.85	–1.03	–1.39	–2.05
0.90	–1.08	–1.44	–2.11
0.95	–1.15	–1.51	–2.18
1.00	–1.28	–1.65	–2.33

Null hypothesis will be rejected if simultaneously

$$u = \frac{U-\overline{U}}{\sigma_U} < h, \qquad v = \frac{V-\overline{V}}{\sigma_V} < h.$$

Example: We suppose x is stochastically smallest and observe for $l = 5$, $m = 5$, $n = 5$ the result

$$x\,x\,x\,y\,x\,y\,z\,y\,x\,y\,z\,y\,z\,z\,z$$

$$U = 4,\ V = 1,\ \overline{U} = 12.5,\ \overline{V} = 12.5,$$

$$\sigma_U = 4.8,\ \sigma_V = 4.8,\ \rho_{UV} = 0.45.$$

From Table X we read that for a correlation of 0.45 the critical value of h at the 5% level is —1.07

$$u = \frac{4 - 12.5}{4.8} = -1.77, \quad v = \frac{1 - 12.5}{4.8} = -2.40.$$

Because both u and v are less (more negative) than the critical value of h, we reject the null hypothesis, and accept x as stochastically smallest.

We turn now to testing whether the three populations are in an order prescribed in advance. It will be a great convenience if we can keep the same definitions for U and V and rearrange the stochastic order. We still have l x's, m y's, and n z's. Let us suppose that the order specified in advance is y smallest, x next smallest, z largest. We order the observations from least to greatest as before. If y is to be smallest, it should appear often before x, and therefore U should be large. If z is to be largest it should seldom appear before x and therefore V should be small, as before. All computations are as before, but we want

$$u \geqslant h \text{ and } v \leqq -h$$

simultaneously, where h is positive.

Example: We suppose y is smallest, x next, z largest, and we observe the order

$$y\,y\,y\,x\,z\,y\,x\,y\,x\,z\,x\,x\,z\,z\,z \qquad l = m = n = 5$$

$$U = 22,\ V = 6,\ \overline{U} = \overline{V} = 12.5,$$

$$\sigma_U = \sigma_V = 4.8, \qquad \rho_{UV} = 0.45,$$

$$u = 1.98,\ v = -1.35.$$

From Table XI we find the critical value of h for the 5% level to be 0.45, and because u is larger than 0.45 and v less than —0.45 we reject the null hypothesis and accept the supposed order.

The problem of m rankings. If we have a two-way analysis of variance table where several judges have each ranked the same set of stimuli, we may wish to measure the concordance (or discordance) of the rankings. Each judge will have used the same ranks, so the total ranking for any one judge is the same. We cannot tell anything, therefore, about differences between judges. We can tell whether the judges are agreeing better than a dummy who does his rankings by random placement. (The same procedure may be used on measurement data if they are appropriately ranked.) For example, Guilford (1936, p. 247) gives ranks assigned to 12 handwriting samples by various judges. We shall use his first 5 judges to illustrate.

TABLE XI

VALUES FOR h FOR TESTING NULL HYPOTHESIS AGAINST ALTERNATIVE HYPOTHESIS

$$y < x < z$$

Correlation	*Significance level*		
r	0.10	0.05	0.01
0.00	0.48	0.76	1.28
0.05	0.45	0.73	1.24
0.10	0.42	0.69	1.20
0.15	0.39	0.66	1.15
0.20	0.37	0.63	1.10
0.25	0.34	0.59	1.06
0.30	0.31	0.56	1.01
0.35	0.28	0.52	0.96
0.40	0.25	0.49	0.91
0.45	0.22	0.45	0.86
0.50	0.19	0.41	0.81
0.55	0.17	0.38	0.76
0.60	0.14	0.34	0.70
0.65	0.10	0.30	0.64
0.70	0.07	0.26	0.58
0.75	0.04	0.21	0.51
0.80	0.01	0.17	0.44
0.85	0.00	0.12	0.37
0.90	0.00	0.07	0.28
0.95	0.00	0.00	0.17

Null hypothesis will be rejected if simultaneously

$$u = \frac{U - \overline{U}}{\sigma_U} > h, \quad v = \frac{V - \overline{V}}{\sigma_V} < -h$$

	Handwriting Sample											
Judge	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*
1	12	8	11	10	9	6	5	4	7	1	2	3
2	11	12	10	9	7	8	6	4	1	3	5	2
3	10	11	12	9	7	8	6	5	3	4	1	2
4	9	12	11	10	8	7	4	3	5	6	2	1
5	12	10	11	9	4	7	8	1	6	2	5	3
Totals	54	53	55	47	35	36	29	17	22	16	15	11

If all stimuli lead to similar outcomes we expect the sum of the ranks for each column to be roughly the same size. But if some stimuli produce larger and others smaller effects, we expect such stimuli to produce differing column sums. To evaluate the significance of the disparity among column sums, we compute χ_r^2, with $p - 1$ degrees of freedom:

$$\chi_r^2 = \left\{ \frac{12}{np\,(p+1)} \times \text{sum (rank total)}^2 \right\} - 3n\,(p+1),$$

where we are summing the squares of column totals, and

n = number of replications (rows),

p = number of treatments (columns).

The subscript r on χ^2 serves notice that this is a χ^2 derived from ranks and, in fact, is an approximation to χ^2 rather than an exact χ^2 statistic.

For Guilford's handwriting example, we get

$$\text{sum (rank total)}^2 = 15{,}696,$$

$$\chi_r^2 = \frac{12\,(15{,}696)}{5\,(12)\,(13)} - 3\,(5)\,(13)$$

$$= 46.5,$$

with 11 degrees of freedom. Clearly this result is highly significant, and it means that the judges agree that the stimuli are different. In a case like the present one where the ranks are obviously in good agreement, the test gives little information that could not be had from casual observation of the table. Although the matter has not had theoretical investigation, we would suggest that the Link-Wallace short-cut to allowances in one-way tables, discussed above, might reasonably be applied here. Rather than testing for concordance, we would be attempting to decide which pairs of stimuli should be "recognized" as different. The method of m rankings was developed by Friedman (1937), but expositions of this and other ranking methods together with references are probably more readily available in Kendall (1943, Chapter 16) and Kendall (1948).

The complete ranked one-way analysis of variance. The previous subsection treated ranked data where the ranking was done only in single rows, rather than over the entire table. A new development (Kruskal and Wallis, 1952) concerns one-way analysis of variance where the ranking is done for all observations in the table. (The same paper offers a survey of the literature of ranking methods.)

The H-test presented below requires that all observations be ranked together, and that the sum of the ranks be obtained for each sample. The test statistic to be computed if there are no ties (no two original observations equal) is

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{C} \frac{R_i^2}{n_i} - 3\,(N+1), \qquad \text{(no ties)}$$

C = number of groups,

n_i = number of observations in the ith group,

$N = \Sigma n_i$, the total number of observations,

R_i = sum of the ranks in the ith group.

Large values of H lead to rejection of the hypothesis of equality of the original C populations. If the n_i are not too small, H is approximately distributed like chi-square with $C - 1$ degrees of freedom. We present tables for sets of three samples whose sizes are each less than six.

If there are ties, each observation is given the mean of the ranks for which it is tied. Then H is computed as above, and divided by

$$\text{correction for ties} = 1 - \frac{\Sigma T}{N^3 - N},$$

where $T = (t-1)\,t\,(t+1) = t^3 - t$ for each set of ties, t is the number of tied observations in a set, and the summation is over all sets of ties. The following table helps:

t	1	2	3	4	5	6	7	8	9	10
T	0	6	24	60	120	210	336	504	720	990

Example (with ties). In Table XII we analyze some artificial data constructed to display the problem of ties. We have three groups of unequal sizes, so $C = 3$.

TABLE XII

EXAMPLE OF THE H STATISTIC FOR RANKED ANALYSIS OF VARIANCE

Treatment					
1		2		3	
Observation	Rank	Obserservation	Rank	Obserservation	Rank
0	3	6	14	2	10
3	12	5	13	8	15
0	3	2	10	0	3
1	7	1	7	2	10
12	16	0	3	0	3
15	17	1	7		
88	18				
	$R_1 = 76$		$R_2 = 54$		$R_3 = 41$
	$n_1 = 7$		$n_2 = 6$		$n_3 = 5$

Ties: 5 at 0, 3 at 1, 3 at 2.

$$\Sigma T = 120 + 24 + 24 = 168.$$

$$N^3 - N = 18^3 - 18 = 5{,}814.$$

$$H = \frac{\dfrac{12}{18\,(19)}\left[\dfrac{76^2}{7} + \dfrac{54^2}{6} + \dfrac{41^2}{5}\right] - 3\,(19)}{1 - \dfrac{168}{5{,}814}}$$

$$= \frac{0.8015}{0.9711} = 0.825.$$

$P(\chi^2 > 0.825) = 0.66$ (2 degrees of freedom).

We conclude that there is not a significant difference between the groups. Note that the huge observation 88 in Treatment 1 ultimately contributed little more than the observation 15. Note also that the correction for ties changed matters very little (about 3%).

Wald-Wolfowitz run test. The run test inquires whether two samples (not necessarily of equal size) are drawn from a common population. No matter in what way populations differ — medians, variability, skewness — the run test will reject the null hypothesis of no difference with as large a probability as we wish with big enough samples.

To make the test, pool the samples and arrange the data in order of increasing size. Moses presents the following example:

$$\underline{X_1X_2}Y_1\underline{X_3}Y_2Y_3Y_4Y_5\underline{X_4X_5}Y_6$$
$$\underline{X_6X_7X_8}Y_7\underline{X_9X_{10}}.$$

TABLE XIII

TRUE DISTRIBUTION OF H FOR THREE SAMPLES, EACH OF SIZE FIVE OR LESS, IN THE NEIGHBORHOOD OF THE 10, 5, AND 1 PERCENT POINTS

The probabilities shown are the probabilities, under the null hypothesis, that H will equal or exceed the values in the column headed H.

Sample sizes			*H*	*True probability*
n_1	n_2	n_3		
2	2	2	4.57	.067
3	2	1	4.28	.100
3	2	2	4.71	.048
			4.50	.067
3	3	1	5.14	.043
			4.57	.100
3	3	2	5.36	.032
			4.55	.100
3	3	3	7.20	.004
			5.60	.050
			4.62	.100
4	2	1	4.50	.076
4	2	2	5.33	.033
			4.45	.100
4	3	1	5.20	.050
			4.05	.093
4	3	2	6.44	.008
			5.44	.046
			4.51	.098
4	3	3	6.74	.010
			5.72	.050
			4.70	.092
4	4	1	6.66	.010
			4.96	.048
			4.16	.082
4	4	2	7.03	.006
			5.45	.046
			4.55	.098
4	4	3	7.14	.010
			5.59	.049
			4.54	.099
4	4	4	7.65	.008
			5.69	.049
			4.65	.097
5	2	1	5.00	.048
			4.20	.095
5	2	2	6.53	.008
			5.16	.034
			4.37	.090
5	3	1	4.96	.048
			4.01	.095
5	3	2	6.82	.010
			5.25	.049
			4.65	.091
5	3	3	7.07	.009
			5.64	.049
			4.53	.097
5	4	1	6.95	.008
			4.98	.044
			3.98	.098
5	4	2	7.11	.010
			5.26	.050
			4.54	.098
5	4	3	7.44	.010
			5.63	.050
			4.54	.099
5	4	4	7.76	.009
			5.61	.050
			4.61	.100
5	5	1	7.30	.009
			5.12	.046
			4.10	.086
5	5	2	7.26	.010
			5.33	.047
			4.50	.100
5	5	3	7.54	.010
			5.70	.046
			4.54	.100
5	5	4	7.79	.010
			5.64	.050
			4.52	.099
5	5	5	7.98	.010
			5.78	.049
			4.56	.100

A run is a group of consecutive values drawn from only one population. Runs of X's are underlined, and runs of Y's appear between the underlined X's. In this example there are nine runs altogether. If two samples are drawn from the same population, we expect that the number of runs will be fairly large, since we expect that X and Y values will be well intermingled. There is no reason to expect all the Y's to fall at one end and all the X's to fall at the other end, for instance, unless the populations are really very different. This situation would lead to our observing only two runs. Too few runs, then, will cause rejection of the null hypothesis.

We present in Table XIV critical values of the number of runs d for m less than or equal to n. Here m and n are the numbers of elements or observations of each kind. This table is used to test at the 5% level whether there are too few runs. If there are too few runs, the populations from which the samples are drawn differ. In our example there are 10 X's and 7 Y's; therefore, $m = 7$, $n = 10$. We observed $d = 9$. Reading from Table XIV, we see that the critical number of runs is 5, and so our result does not imply too few runs.

A somewhat different problem is that of testing randomness of two kinds of elements. For example, a half-hour before the popular lunch hour the following arrangement of empty and occupied seats at a lunch counter might be observed (Swed and Eisenhart, 1943):

$$E\,O\,E\,E\,O\,E\,O\,E\,O\,E\,E\,E\,O\,E\,O\,E$$

The question is whether the people are randomly seated or whether they prefer elbowroom (leading to too many runs). Here we have 10 E's and 6 O's, so $m = 6$ and $n = 10$; also the number of runs is $d = 13$. From Table XV, the critical value of d is 12. Because our observed number is larger, we reject the hypothesis of random seating.

For values of m and n outside the table we can use as a normally distributed variable (Mood, 1950, p. 393)

$$z = \frac{d - \bar{d} \pm \frac{1}{2}}{\sigma_d},$$

where

$$\bar{d} = \frac{2mn}{m+n} + 1,$$

$$\sigma_d{}^2 = \frac{2mn\,(2mn - n - m)}{(m+n)^2\,(m+n-1)}.$$

In the correction for continuity we use $+\frac{1}{2}$ if we are interested in the probability of d or fewer runs, $-\frac{1}{2}$ if we want the probability of d or more runs.

TABLE XIV

CRITICAL VALUES OF d (NUMBER OF RUNS) FOR VARIOUS VALUES OF m AND n, THE NUMBERS OF THE TWO KINDS OF ELEMENTS

The values tabled are the largest values of d that give significance at the 5% level. If the observed d is as small as or smaller than the entry, there are too few runs.

$n \backslash m$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	—																		
3	—	—																	
4	—	—	2																
5	—	2	2	3															
6	—	2	3	3	3														
7	—	2	3	3	4	4													
8	2	2	3	3	4	4	5												
9	2	2	3	4	4	5	5	6											
10	2	3	3	4	5	5	6	6	6										
11	2	3	3	4	5	5	6	6	7	7									
12	2	3	4	4	5	6	6	7	7	8	8								
13	2	3	4	4	5	6	6	7	8	8	9	9							
14	2	3	4	5	5	6	7	7	8	8	9	9	10						
15	2	3	4	5	6	6	7	8	8	9	9	10	10	11					
16	2	3	4	5	6	6	7	8	8	9	10	10	11	11	11				
17	2	3	4	5	6	7	7	8	9	9	10	10	11	11	12	12			
18	2	3	4	5	6	7	8	8	9	10	10	11	11	12	12	13	13		
19	2	3	4	5	6	7	8	8	9	10	10	11	12	12	13	13	14	14	
20	2	3	4	5	6	7	8	9	9	10	11	11	12	12	13	13	14	14	15

Moses points out that tests of this kind which guard against many alternatives are not very good at guarding against some *one particular* alternative. Thus, to test specifically whether two samples came from a population with the same median, we will be more likely to detect this departure using the Mann-Whitney test, a test specifically designed to guard against the alternative of interest. Furthermore, when a test guards against many alternatives and we reject the null hypothesis, we find ourselves in the unsatisfactory position of not being able to pin down what alternative (s) has led to the rejection. To many persons engaged in experimental work it is often the specific nature of a departure that is interesting, but, as with many other things, it is only by sacrificing some generality in the hypothesis that particular specificities in the data have a chance to display themselves.

Table XV

CRITICAL VALUES OF d (NUMBER OF RUNS) FOR VARIOUS VALUES OF m AND n, THE NUMBERS OF THE TWO KINDS OF ELEMENTS

The values tabled are the smallest values of d that give significance at the 5% level. If the observed d is as large or larger than the entry, there are too many runs.

$n \setminus m$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	—																	
4	7	8																
5	—	9	9															
6	—	9	10	11														
7	—	9	10	11	12													
8	—	—	11	12	13	13												
9	—	—	11	12	13	14	14											
10	—	—	11	12	13	14	15	16										
11	—	—	—	13	14	15	15	16	17									
12	—	—	—	13	14	15	16	17	17	18								
13	—	—	—	13	14	15	16	17	18	18	19							
14	—	—	—	13	14	16	17	18	18	19	20	20						
15	—	—	—	—	15	16	17	18	19	19	20	21	21					
16	—	—	—	—	15	16	17	18	19	20	21	21	22	23				
17	—	—	—	—	15	16	17	18	19	20	21	22	22	23	24			
18	—	—	—	—	15	16	18	19	20	21	21	22	23	24	24	25		
19	—	—	—	—	15	16	18	19	20	21	22	23	23	24	25	25	26	
20	—	—	—	—	—	17	18	19	20	21	22	23	24	25	25	26	27	28

Quick Methods

Standard deviation of a mean. Often one wants to know quickly an approximate estimate of the standard deviation of a mean for a small set of data (in order to make a rough *t*-test, for example). A good approximation (Mantel, 1951) can be obtained quickly by dividing the range by the total number of observations, i.e.,

$$\frac{\text{largest observation} - \text{smallest observation}}{\text{number of observations}}$$
$$= \text{estimated standard deviation of mean}$$

for samples of size 15 or less. This estimate seems to be good regardless of the form of the population distribution — the estimate works well for even the uniform distribution up to at least 10. For example, with the six observations 6, 7, 10, 12, 15, 18 Mantel's procedure gives $\frac{18-6}{6} = 2$. The longer root mean square formula gives 1.89.

t-Tests. To test whether a mean is equal to a standard or whether two samples are drawn from populations with equal means, Lord (1947) suggests two tests based on the range suitable for samples up to about size 20. He suggests replacing the ordinary *t*-test by the statistic

$$\frac{\bar{x} - m}{R},$$

where $\bar{x}$ is the sample mean and R is the range of the sample. Similarly, in testing for the equality of means with two samples of equal size he proposes the statistic

$$\frac{\bar{x}_1 - \bar{x}_2}{\frac{1}{2}(R_1 + R_2)},$$

where the $\bar{x}$'s are the sample means and the R's represent the ranges of the samples. Both tests are for use with small samples, say samples of sizes from 2 to 20. Lord provides special tables for using these statistics in significance testing. Since these tables are not widely available and since it is possible to relate them closely to a t-table by introducing a conversion factor, we give conversion factors which will make it possible to use these statistics with an ordinary t-table and with the same numbers of degrees of freedom as are customarily used in such tests. In particular, for the one-sample case

$$t_{n-1} = \left(\frac{\bar{x} - m}{R}\right) C_n,$$

where

$$C_n \doteq 5\sqrt{n + 0.5} - 6.$$

The t_{n-1} can be used for the one-sample test, where the subscript $n - 1$ on the t refers to the number of degrees of freedom for entering an ordinary t-table.

In the two-sample case,

$$t_{2n-2} = \left(\frac{\bar{x}_1 - \bar{x}_2}{\frac{1}{2}(R_1 + R_2)}\right) K_n,$$

where

$$K_n \doteq 4\sqrt{n + 2.4} - 7.14.$$

In Table XVI we present values of the correction terms C_n and K_n for single samples of size $n = 2$ to 20. Thus the approximate formulas need not be used. It should be understood that if one uses C_n and K_n, either from the tabled values or the formulas, one will not get exactly the results given by Lord's tables, but they will be extremely close.

The principal value of such statistics is that they avoid the necessity of the root mean square computations ordinarily used, while sacrificing little accuracy.

TABLE XVI

FACTORS FOR CONVERTING LORD'S ONE-SAMPLE AND TWO-SAMPLE t-TESTS TO ORDINARY t-TABLES

Here n is the sample size. We multiply $\dfrac{\bar{x} - m}{R}$ by C_n and enter a t-table with $n - 1$ degrees of freedom. For two samples, both size n, we multiply $\dfrac{\bar{x}_1 - \bar{x}_2}{\frac{1}{2}(R_1 + R_2)}$ by K_n and enter an ordinary t-table with $2n - 2$ degrees of freedom.

Single sample size n	*One-sample test* C_n	*Two-sample test* K_n
2	2.00	1.26
3	3.30	2.19
4	4.44	3.02
5	5.48	3.77
6	6.44	4.47
7	7.35	5.14
8	8.21	5.75
9	9.04	6.35
10	9.83	6.94
11	10.6	7.50
12	11.3	8.02
13	12.0	8.51
14	12.7	9.03
15	13.4	9.50
16	14.1	9.98
17	17.7	10.5
18	15.4	10.9
19	16.0	11.3
20	16.6	11.8

TRANSFORMATIONS

Any method of systematically transforming a given set of numbers into a new set may be regarded as a transformation. Such a general definition does not get us very far, so let us discuss a few examples. (i) We are familiar with transforming observations to standard scores by multiplying each observation by one number and then adding another number. This transformation gives us a new set of observations whose mean and standard deviation have been adjusted to our satisfaction. If individuals have been measured on several characteristics, and each measure expressed in standard scores, then an individual has a profile of scores readily interpretable at a glance. One use, then, of transformations is to make data more comparable. (ii) The standard score transformation is just an example of the general linear transformation $y = a + bx$, where x is an original observation and a and b are constant. This general transformation makes it possible to adjust a set of numbers so they have any desired zero point and any desired unit of measurement. (iii) Another transformation that often improves comparability is the use of percentages or ratios rather than raw numbers.

This use is so common we shall not discuss it further. (iv) We often see graphs of one variable against another where the relation between the two is a straight line, but one axis represents, say, the logarithm of the time, and the other perhaps rate of responding. Straight lines always look good in graphs, and most authors prefer to mangle (transform) one or both of the variables if there is any possibility of displaying linearity. Such adjustments sometimes suggest theoretical possibilities. (v) To get back to analytical statistics, we have seen transformations to ranks used on a number of occasions in the present chapter. We speak of not needing to know the original form of the distribution if we use ranks. Just what does this mean? Take the Mann-Whitney test as an example. If we do indeed have two identical continuous populations and we draw a sample of two observations from each, then we know that each of the six possible outcomes of ordering the four observations has exactly the same chance of occurring. They are:

$x\,x\,y\,y$	$y\,x\,x\,y$
$x\,y\,x\,y$	$y\,x\,y\,x$
$x\,y\,y\,x$	$y\,y\,x\,x$

The simple process of ranking the observations then replaced a situation where we knew practically nothing usable about the original observations by a new situation where we know everything about the distribution of the ranks when the null hypothesis is true. Thus the transformation to ranks can sometimes divorce us from the unknown. (vi) The transformation to signs plays the same role for the sign test. (vii) When we replace an observed value of chi-square or of a normal deviate by a probability level read from a table we are making a transformation whose utility lies in its ease of interpretation as a frequency of occurrence. (viii) Some people make transformations so that their adjusted data will be more nearly normally distributed than the original observations, for example, the logarithms of the lengths of sentences written by some famous authors are approximately normally distributed, whereas the original distribution of number of words has a rather extended tail and is quite asymmetric. In this example, a good deal of information about the original distribution can be transmitted by reporting the mean and standard deviation of the logarithmic sentence lengths. Such transformations designed to achieve approximate normality are not alone a feature of empirical distributions but also are used for the theoretical distributions. To illustrate, many chi-square tables carry instructions in a footnote to the effect that beyond $n = 30$ degrees of freedom the expression $\sqrt{2\chi^2} - \sqrt{2n - 1}$ is approximately a normal deviate with zero mean and unit standard deviation. Similarly, the z-transformation for testing the significance of correlation coefficients is a device for achieving approximate normality.

We wish to describe a type of transformation that has a slightly different purpose from the ones mentioned above. In many kinds of measurements the larger the thing being measured the more variable the measurements. For example, ask subjects to close their eyes and press a button 2 seconds after a buzzer sounds, and then ask them to press after 60 seconds. Times till pressing will be much more variable for the 60-second data than for the 2-second data. If we have data for several such times we will find then that the different times do not have the same variability. (This terrible statistical disease has been christened "heteroscedasticity," a fact which has blighted the lives of many students. Some years ago much was made of other statistical diseases such as kurtosis, but the fad seems to have passed.) Sometimes a transformation such as taking logarithms, square roots, or cube roots of the original data will make the variances of the adjusted observations approximately equal for each time or, more generally, each condition of an experiment. If such a transformation can be found it is called, appropriately enough, a "variance-stabilizing transformation." An important feature of a variance-stabilizing transformation is that it leaves the group means in a reasonable order. A few readers may have been thinking "Why not convert each time group to standard scores?" But this technique will remove the relation between instruction time and observed waiting time, which is of interest to the experimenter. Logarithms or square roots will at least leave the increase of observed time definitely dependent on the length of time requested. But why do we want equal variances anyhow? One of the requirements of analysis of variance is that each cell of the original table should represent a normal population with the same variance, although not necessarily the same mean. So a move to transformed data with more nearly equal variances would more nearly meet the requirements of analysis of variance. To suggest a specific use, suppose we try to use the residual mean square to assist in setting confidence limits on the mean of a specific cell in the analysis of variance table. If we happen to have unequal variances in the cells we may find that the cell we chose has observations in it that are outrageously disparate compared with the prediction that is being made. Had we made a variance-stabilizing transformation before the analysis, we could have set confidence limits on the transformed mean, and then transformed these limits back to the original units without the same risk of error.

Up to this point we have talked about variance-stabilizing transformations in the absence of theoretical notions about the form of the distribution. Much data is derived from counting things. If we have a fixed number of trials, we count the number of successes. If a subject has the same chance of success p on each trial, then scores for this subject should be binomially distributed with mean proportions of successes p, and variance of the proportion from one set of trials to another $p\,(1 - p)\,/n$, where n is the number of trials in a set. Note that the variance depends on the mean. When this happens we can usually find an approximate variance-stabilizing transformation, and for the binomial we shall give one below. We shall also give an additional reason for the use of such a transformation.

A second kind of counting distribution is the Poisson. It usually arises when we have a fixed exposure (in space or time or both) and count the number of events occurring in the exposure. For example, subjects observe a staircase illusion for 20 seconds and press a button every time it changes position. If the subject does this several times under nearly identical conditions we expect the distribution of counts to be approximately Poisson. If the mean number of alternations per occasion is m, then the variance is also m. Again we can find an approximate variance-stabilizing transformation. It turns out to be the square-root transformation. To take another example of a distribution that may be approximately Poisson, Eichler (1951) proposes using the square-root transformation on Rorschach scores as a method of reducing the influence of total productivity on subscores. Because the total number of productions of a particular Rorschach sign is small, it would be reasonable to suppose the distribution of the number on repeated takes by the same person to be approximately Poisson. The square-root transformation is appropriate for Poisson data. Counts from binomial or Poisson data (or Rorschach data) have variances that depend on the unknown true means — this difficulty can often be circumvented by using the inverse sine transformation and the square-root transformation to obtain new observations whose variances do not depend much on their means.

Freeman and Tukey (1950) suggest for the Poisson that the variance-stabilizing properties of the square-root transformation are improved by using

$$\sqrt{x} + \sqrt{x+1},$$

where x is the observed count, instead of just $\sqrt{x}$ or $\sqrt{x + \tfrac{1}{2}}$. The variance of Poisson data so transformed is approximately unity. In Table XVII we provide a set of tables of this square-root transformation for Poisson variables and its square for x running from 0 to 50. This table should facilitate making analysis of variance computations using this transformation.

The third column in Table XVII is the square of the two-decimal place number given in the second column, rather than the correct four-decimal result, because these are the appropriate squares to use if the two-decimal transformation in the second column is employed for analysis of variance purposes.

For the binomial Freeman and Tukey propose a slight modification of the usual transformation, i.e.,

$$\tfrac{1}{2}\left(\sin^{-1}\sqrt{\frac{x}{n+1}} + \sin^{-1}\sqrt{\frac{x+1}{n+1}}\right),$$

and data thus transformed have approximately variance $\dfrac{821}{n+0.5}$ for angles in degrees. Here x is the count, n is the sample size. Values of $\sin^{-1}\sqrt{p}$ can readily be obtained because Snedecor (1946, pp. 449-450) has reproduced C. I. Bliss's tables of this function.

The inverse sine transformation must be computed separately for each sample size. We have investigated how well the Freeman-Tukey inverse sine transformation stabilizes the variance, and below the reader can see that this transformation does indeed produce values very close to the theoretical value of $\dfrac{821}{n+0.5}$ even for p's as low as 0.10 and small sample sizes. We show results of two computations, one for $n = 10$ and one for $n = 20$.

	Theoretical variance		*Computed variance*				
n		$p=0.50$	$p=0.40$	$p=0.30$	$p=0.20$	$p=0.10$	$p=0.05$
10	78.2	76.1	76.7	78.8	81.8	74.9	53.7
Ratio of computed to theoretical		0.97	0.98	1.01	1.05	0.96	0.69
20	40.0	39.2	39.2	39.3	40.0	42.2	38.0
Ratio of computed to theoretical		0.98	0.98	0.98	1.00	1.05	0.95

TABLE XVII

TABLE OF VALUES FOR THE FREEMAN-TUKEY SQUARE-ROOT TRANSFORMATION

The observed count is x and the transformed values are $\sqrt{x} + \sqrt{x+1}$. The squares of the transformed values are also given for use in analysis of variance computations.

x	$\sqrt{x}+\sqrt{x+1}$	$(\sqrt{x}+\sqrt{x+1})^2$	x	$\sqrt{x}+\sqrt{x+1}$	$(\sqrt{x}+\sqrt{x+1})^2$
0	1.00	1.0000	26	10.30	106.0900
1	2.41	5.8081	27	10.49	110.0401
2	3.15	9.9225	28	10.68	114.0624
3	3.73	13.9129	29	10.86	117.9396
4	4.24	17.9776	30	11.04	121.8816
5	4.69	21.9961	31	11.22	125.8884
6	5.10	26.0100	32	11.40	129.9600
7	5.47	29.9209	33	11.58	134.0964
8	5.83	33.9889	34	11.75	138.0625
9	6.16	37.9456	35	11.92	142.0864
10	6.48	41.9904	36	12.08	145.9264
11	6.78	45.9684	37	12.25	150.0625
12	7.07	49.9849	38	12.41	154.0081
13	7.35	54.0225	39	12.57	158.0049
14	7.61	57.9121	40	12.73	162.0529
15	7.87	61.9369	41	12.88	165.8944
16	8.12	65.9344	42	13.04	170.0416
17	8.37	70.0569	43	13.19	173.9761
18	8.60	73.9600	44	13.34	177.9556
19	8.83	77.9689	45	13.49	181.9801
20	9.05	81.9025	46	13.64	186.0496
21	9.27	85.9329	47	13.78	189.8884
22	9.49	90.0601	48	13.93	194.0449
23	9.69	93.8961	49	14.07	197.9649
24	9.90	98.0100	50	14.21	201.9241
25	10.10	102.0100			

As one would expect, the stability of the variance improves as the sample size increases. Freeman and Tukey show that their square-root transformation stabilizes Poisson variance satisfactorily for true means as small as 1.00.

The most valuable property of these transformations does not seem to be the variance-stabilizing property, however. Rather, when these transformations are used we are provided with a *known base-line* variance. In an analysis of variance of the transformed data the error variance can be compared with the theoretical error variance. This provides an experimenter with important knowledge, for if he can estimate how much of the variance in an experimental situation is "natural variance," he is in a good position to assess the residual variance. Thus, he can decide whether to try to improve control and, if he does, he has a measure of how much the new control reduces the error variance. Hence, for the practical man, these transformations give not only better statistical tests, but also a notion of the limiting accuracy of the experiment or, alternatively, they give a measure of the variability in an experiment owing to causes that could in principle be controlled.

To give an example or two, different subjects have studied a long word list 0, 10, 30 minutes. Words from the list are exposed in a tachistoscope at three different exposure times. A different subject observes 20 words for each condition of training and each exposure time. The number of correct recognitions is recorded. The distribution of successes may be binomial. If the inverse sine transformation is used and an analysis of variance performed we anticipate the residual mean square (error variance) to be $821/20.5 = 40.0$. If we observe a residual mean square of from 40 to about 80 (depending on the number of degrees of freedom), we will be fairly satisfied with the control achieved. Suppose we observe 160. We begin to wonder whether there are not factors in the experiment that ought to be further controlled. Perhaps the subjects vary in ability to study the list and in visual acuity, and it would be better to use the same subject for all conditions, so that he serves as his own control. Perhaps a replication within each cell would provide a better estimate of the base-line error, because any interaction between exposure time and amount of training would contaminate an error time. What if, far from observing a mean square of 40, we observe a small number like 10? First, look for a mistake in computation. If one cannot be found by you or an independent computer, reconsider the experiment. Such a thing might happen in the present experiment (if it were poorly designed) if the same list of 20 words were always used, and a certain few were always recognized and another set were never recognized. Thus the actual variation in number recognized would be less than anticipated because the words on the list were not equated for ease of recognition. However, it is a rather rare day when the observed residual variation is smaller than anticipated.

As another example, nursery school children are observed during the second 10 minutes of a regular play hour for several days at the beginning, middle, and end of their first semester. There are several types of play programs used throughout the year. The observers record the number of acts of overt aggression against others made by each child during these recording periods. Here the number of such acts might be approximately Poisson distributed, especially if an act is defined so that an extended altercation by the same pair of children only counts once. So it might be reasonable to try the square root transformation, in which case the base-line variance is 1. If the observed residual variance were a good deal larger than this we would look around for matters other than time of year, program, or between children differences to help account for the extra variability. In this example, variation of scoring procedure and unreliability of scoring are important possible contributors to variance. On the side of small variance there is the occasional retiring child who commits practically no acts in the category under study.

Two warnings are in order. First, the fact that data appear in percentage form does not automatically justify the assumption of binomiality, and therefore the inverse sine transformation is not always appropriate. As examples: the percentage of body weight consumed at a sitting under free feeding, or the percent of a color wheel made blue to match a stimulus card are not cases where the variance-stabilizing properties would be expected on the basis of binomial theory. Secondly, when an analysis of variance is planned for binomial data, it is well to try to assure in advance that the sample sizes in the cells are always the same. If there are unequal numbers of opportunities for success in the cells a very sticky problem arises which we cannot treat here, but see Cochran (1943, pp. 287-301).

COMBINING TESTS OF SIGNIFICANCE

Suppose an experimenter wants to combine the results of tests made on several independent groups of data into a single grand significance test for the set. "There may be several samples which cannot or should not be combined, perhaps because reported by different investigators or because the original data are lacking, but to each of which separately a test of significance can be applied" (Wallis, 1942, p. 229). Before discussing the principal techniques for combining, we wish to forewarn the reader of the importance of independence of the separate results. Suppose a pollster asks a sample of men and women several different questions, but each question is more or less related. The pollster thinks the women's responses show more hypochondriasis than the men's on several of the questions. The pollster wants to pool results from several items to show the over-all significance of a sex difference. The methods of combining significance tests to be discussed here are *inappropriate,* because the several different measures (i.e., responses), almost certainly correlated, are being combined. If the methods of this section are applied and "no significant difference" is concluded, fine. But if "significant difference" is concluded, all that is implied is that probably the answers to the several questions are correlated and we can say nothing about a man-woman difference. This pitfall is charted early and at length, because it has taken a large toll of the unwary in the past. Indeed, one of the main functions of this section is to warn of the misapplication of combining dependent results.

We can combine significance tests when we have independence of the separate tests; an example is when we have test results for several groups of individuals (all different) in the same or possibly different laboratories. The results of the single-group analyses vary. We want to pool the results and make a single significance test. Two examples where combination is appropriate are correlation between frustration and aggression for each of several groups of different subjects, and *t*-tests between scores of rural and urban children on a personality test. In fact, almost any statistic which leads to a probability statement may be combined if it meets the criterion of independence.

First, a word about a method that almost anyone's intuition suggests — just multiplying the observed values of P together and calling the result the significance level. Suppose we agree to reject if the product of the probabilities from two experiments is 0.05 or less, how often will we reject when the null hypothesis is true? The answer, remarkably enough, is 0.20, and not 0.05, as common sense might have suggested. To see why, we need an example. Suppose $P_1 = 0.5$ and $P_2 = 0.1$; their product is 0.05 and we would reject by the above rule. It is perfectly true that the probability of points with $P_1 \leq 0.5$ and $P_2 \leq 0.1$ is 0.05. But we would also reject if we had $P_1 = 0.1$ and $P_2 = 0.5$, so there are additional points for which we would reject. The total probabilities of all these points adds up to 0.20 — a good distance from the desired 0.05. This intuitive procedure would be fine if we had easily available the probability distribution of such products in tabular form. Unfortunately we do not, but more unfortunately persons who want to do this are apt to overlook the fact that the distribution of products of even two probabilities is far from uniform. Since the appropriate distribution is not available, people who combine tests this way are not in a position to evaluate the likelihood of their observed result. The Fisher method described below properly takes care of this.

The fundamental property back of most combining methods is that every probability value is equally likely if the null hypothesis is true. This says that there are just as many outcomes with P between 0.1 and 0.2 as there are between 0.7 and 0.8 (or any other interval of the same size). This implies that any transformation we can find that makes it easy to add up several results is all we need to be ready to combine tests. "Easy to add up" means that whatever final statistics we come out with, it should be easy to compute and have a distribution that is widely available in tables.

Fisher (1938, pp. 104-106) noticed that if P is an observed probability level, then

$$-2 \log_e P$$

has the chi-square distribution with two degrees of freedom. Now the sum of quantities *independently* distributed according to chi-square has again the chi-square distribution with degrees of freedom equal to the sum of the degrees of freedom of each of the component chi-squares. So Fisher's proposal is to combine significance tests by taking minus twice the sum of the natural logarithms of the probabilities. We are thus able to go to a chi-square table to complete the test. The method has some intuitive justification because when we add logarithms, we are performing the same operation as multiplying numbers, as discussed above. Gordon *et al.* (1952) have provided an extended table of chi-square for two degrees of freedom which should facilitate the computations involved in combining probabilities from independent samples. The reason for the table is that one wants to work with values of P which are correct to two or three decimals and to have chi-square values that are correspondingly accurate. Furthermore, their table makes it possible to go directly to the chi-square values and sum these and thus avoid the intervening step of using logarithms altogether. Such a table is a great help in two ways: we avoid the interpolation problem and we avoid the logarithm problem.

Others, see for example Stouffer *et al.* (1949, p. 45, footnote 15), have suggested that we transform the observed P-value to a standard (unit) normal deviate, using any of the standard normal tables that are available. The sum of normal deviates is again a normal deviate with variance equal to the number of observations summed. So the final statistic we refer to the normal table is

$$z = \frac{\text{sum of normal deviates}}{\sqrt{\text{number of observations}}},$$

and we note in passing that it has the nice additive property we want, and also is a statistic widely tabulated. Stouffer's method is easier to compute than Fisher's, because starting from the same basic data all we need for Stouffer's method is a standard normal table and the computations are minimal.

One peculiarity of the Fisher approach may have occurred to the reader. It happens that a large set of P's, all less than 0.5, does not necessarily lead to a significant result — this defies one's intuition about what should happen with a sufficiently large number of results. This result is not encountered with Stouffer's method. If we have many P's, ranging from say 0.37 to 0.49, Stouffer's method would achieve significance as the number of groups increased, ultimately giving a very small (nearly zero) combined probability level. For the same set of P's, the Fisher chi-square values would ultimately lead to a very large combined probability (nearly unity).

A third method of combining significance tests is suggested by a footnote in Gordon *et al.* (1952, p. 315). We think the suggestion flows from analysis of variance considerations. If several laboratories are involved, it may not actually happen, when the null hypothesis is true, that the distribution of *P*-values is uniform. Perhaps there are deviations from uniformity so that there are actually laboratory-to-laboratory differences. If this is the case and we use, say, Stouffer's method, we may be underestimating the actual variance of the normal deviates by assigning their variance the value unity. The result would be that we would too often detect, as real differences, variations owing to laboratory differences rather than to systematic trend. Now, in addition, suppose the null hypothesis is in fact false, then it could easily occur that the standard deviation of the transformed deviates is actually smaller than the value unity assigned, and we might benefit from increased sensitivity by using as a measure of variation the estimated variance from the observed data. The suggestion then would be that one transform from the observed values of *P* to normal deviates, and consider these deviates as a sample from a normal distribution with unknown variance. Then a *t*-test is performed on these deviates to determine whether their population mean could reasonably have been zero, using as degrees of freedom one less than the number of groups. At this writing we would recommend this modified Stouffer technique, with one reservation. When there are very few groups (say 4 or fewer), the size of *t* required for significance is very high. In a sense, too much is being paid for the value of estimating the variance. For this reason most people may prefer the well-established Fisher approach or Stouffer's technique unmodified.

There are three further matters that need serious attention. The first has to do with combining significance tests when the distribution of the original statistic is discrete, for example, combining results from several 2 × 2 tables. The exact significance test results in a discrete distribution. The sign test would be another example, or any of the nonparametric tests (aggravated when sample sizes are small). In such cases the ordinary methods of combining can lead to serious misinterpretations. We cannot go into the matter here, but we refer the reader to Wallis (1942) for a careful discussion.

The second matter has to do with directionality of result. This discussion can best be followed with a *t*-table open before us. Suppose we have three *t*-tests on a measure for men minus a measure for women with *t*-values —0.70, 1.37, 2.23 with 10 degrees of freedom each. What probabilities should we combine? Often people are planning to use a two-tailed test, so they would ordinarily say that the probabilities associated with these three *t*'s are respectively 0.50, 0.20, 0.05. They would take these probabilities and pool them. We feel it is better practice to choose a direction as positive, obtain probabilities of getting the observed value or a larger one, and pool the resulting *P*'s. In the present case we would have as *P*'s 0.75 (because of the minus sign on the original *t*), 0.10, and 0.025. After combining these *P*'s we can use a two-tailed test for the grand result if we so desire. Notice that if we use a two-tailed *P* in the first place we are in the position of suggesting that *t*-values like —2.0 and +2.0 both contribute the same small *P* to the over-all result, thus enhancing its significance. This may cast doubt on the null hypothesis, but the appropriate alternative hypothesis is that these results are too far from one another. The unwary investigator proceeding in the ordinary way might blunder by thinking that he had established a significant difference in a positive direction. Consequently when the original test statistic yields *P*'s corresponding to a two-sided result, we recommend that directionality be established and that it be decided whether an observed two-sided *P* such as 0.10 should be transformed into a one-sided *P* of 0.05 or of 0.95 before proceeding. The chi-square test for large 2 × 2 tables is a good example of the case in point (disregarding the discreteness issue). The results yield a value of chi-square which corresponds to a two-tailed test, directionality may have been confounded, and we may have been measuring heterogeneity of 2 × 2 tables rather than consistency in a particular direction.

A third matter has to do with sample size. In combining *P's* up to now we have spoken as if each experiment to be combined was to be given equal weight. That is the way Fisher originally introduced the test. If we use Fisher's technique, it is not obvious how we can make use of weights, although a theoretical investigation might provide such information. If Stouffer's technique is used (unmodified) there seems to be an obvious way to do the weighting. If we assume that the transformed deviates are independently normally distributed with zero mean and unit variance (null hypothesis), then the statistic

$$T = w_1x_1 + w_2x_2 + \cdots + w_nx_n,$$

where the *w*'s are the positive weights, to be decided upon by the investigator, and the *x*'s are the transformed deviates, will have zero mean and variance

$$\sigma_T^2 = w_1^2 + w_2^2 + \cdots + w_n^2.$$

This means that $z = \dfrac{T}{\sigma_T}$

can be regarded as a normal deviate, and no difficulty arises from the weighting. In the modified Stouffer method using a t-test, it again is not obvious how to proceed with weighting, although perhaps something could be worked out.

We now give an example of the equally weighted methods, treating the same data all three ways. The example is from McNemar and Terman (1936) as given by Gordon *et al.* (1952). Boys and girls were compared on a paper-and-pencil intelligence test in thirteen age groups. We give here the sample number, the one-sided probability of more extreme deviations, the two-degree of freedom chi-square value, and the corresponding normal deviate. In each sample there were about equal numbers of boys and girls, minimum sample size 325, maximum 1,829.

(i) Using Fisher's technique, we look at a chi-square table with 26 degrees of freedom. We find that 43.8 for a one-sided test is between the 0.025 and 0.01 and therefore for a two-sided test it is between 0.05 and 0.02. (ii) Using Stouffer's technique, we find $z = \dfrac{10.14}{\sqrt{13}} = 2.81$, which for a normal table gives a one-sided level of 0.0025, or a two-sided 0.005 level. (iii) Using the t-test technique, we get for 12 degrees of freedom

$$t = \frac{0.78}{0.187} = 4.17,$$

which for a one-sided test gives between 0.005 and 0.0005, therefore a two-sided result between 0.01 and 0.001. It will be noted in the present example that the normal and t-test approach are in close agreement, while the chi-square result is not very close to the others. The fact that in this one case the normal deviate approaches were more sensitive than the chi-square approach should not be construed as evidence in their favor.

	One-sided P	*Chi-square with 2 degrees of freedom*	*Normal deviate*
1	0.159	3.6777	1.00
2	0.100	4.6052	1.28
3	0.021	7.7265	2.04
4	0.386	1.9038	0.29
5	0.405	1.8077	0.24
6	0.054	5.8375	1.61
7	0.111	4.3965	1.22
8	0.371	1.9831	0.33
9	0.309	2.3488	0.50
10	0.255	2.7330	0.66
11	0.121	4.2239	1.17
12	0.648	0.8677	−0.38
13	0.429	1.6926	0.18
		43.8040	10.14
			$\bar{x} = 0.780$

Final Remarks

We have assembled here a collection of technical statistical devices we have found useful in helping to handle problems brought to us by social psychologists, so no great purpose would be served by trying to summarize these somewhat unrelated techniques. It might be better to limit ourselves to a few general remarks about statistics and mathematical devices.

We first wish to emphasize that formal statistics provides the investigator with tools useful in conducting thoughtful research; these tools are not a substitute for either thinking or working. A major goal for the statistical training of students should be statistical thinking rather than statistical formulas, by which we mean specifically: thinking about (1) the conception and design of the study and what it is that is to be measured and why, (2) the definitions of the terms being used, and how modifications in definition might change both the outcome and the interpretation of a study, (3) sources of variation in every part of the study, including such things as individual differences, group and race differences, environmental differences, instrumental or measuring errors, and intrinsic variation fundamental to the process under investigation. In no circumstances do we think that sophisticated analytical devices should replace clean design and careful execution, unless very unusual economic considerations arise. However, it may be worth remarking that crude data collected as best the investigator could may require the most advanced statistical tools. Here a quotation from Wallis (1949) may be appropriate:

> In general, if a statistical investigation . . . is well planned and the data properly collected the interpretation will pretty well take care of itself. So-called "high-powered," "refined," or "elaborate" statistical techniques are generally called for when the data are crude and inadequate — exactly the opposite, if I may be permitted an *obiter dictum,* of what crude and inadequate statisticians usually think.

Next we shall comment on the use of significance tests. It has been aptly said that the main purpose of a significance test is to inhibit the natural enthusiasm of the investigator. The test is used to give some way of deciding whether to believe in the reality of the observed effect. Having established the reality, there is another question—whether the size of the difference is of any material value. In a weight-reducing experiment an average reduction of half a pound can scarcely contribute much to the additional longevity of men who weigh over an eighth of a ton, however real the decrease may be. Such a point has often been made, and it can scarcely be made too often.

Questions often arise about the choice of significance level, but it cannot be expected that any general answer will be forthcoming. The development of the statistical notion of decision functions rather than classical significance tests emphasizes the point all the more. In the decision function approach one tries to decide how much data is needed to reach a decision which in some sense is most economical, but for the full treatment it is important to know the value of correct decisions and the costs associated with incorrect ones. In scientific work, as far as we can tell, we have only the vaguest notions of the costs of mistakes or the value of a correct finding. Often scientific work is not intended as a basis for action; we merely want to know what to believe. How shall we assess the value? What metric shall we use? The answer seems to be "to each his own," and aside from general notions provided by the operating characteristics of various techniques of deciding, there is little we can offer. So the choice of significance level seems to be a matter for the investigator. As for a related topic of one-sided versus two-sided tests, we think there are plenty of situations in which it is appropriate to test for directionality, but a factor of two in the significance level is scarcely worth arguing about. If it is a question of deciding between 0.05 and 0.10, the experiment was not ecstatically conclusive anyway.

We have earlier alluded to the fact that significance levels are, at best, approximate values. Reasons for this are obvious—the assumptions on which the null hypothesis computations are based may be somewhat in error. Independence from trial to trial, for example, may be lacking, randomization may be improperly performed, or the observations may not be normally distributed. Such departures from the ideal may not produce large errors in the true significance level, but they can make it only approximate.

We wish to turn now to the question of other mathematical matters. Except for analysis of variance and multiple regression the social psychologist, as such, has not generally put great efforts into multivariate techniques, such as factor analysis. Without choosing a particular brand of multivariate analysis to push, it still seems reasonable to us that this direction of investigation is one appropriate for the discipline as far as quantitative methods are concerned. An anecdote from the classroom of one of the authors sums up the attitude of a reasonable number of psychologists. In the early part of the course dealing with one-variable problems the complaint was registered by a student: "The problems of social psychology are so much more complicated than this. How can you hope to handle the many variables surrounding a human problem with one variable?" Some weeks later after an introduction to factor analysis in which two factors were removed from 12 variables, the same student said "My God, isn't there some way we can just reduce this all to a single number?" The reply was that essentially multivariate problems are likely to have multivariate answers. We are not arguing here for any special kind of multivariate analysis, but rather for a little more intense investigation of the subject. This is a direction that may pay off, especially if the analysis is built for the problem rather than vice versa. Consequently, social psychologists may wish to persuade more mathematical statisticians to explore the possibilities.

In this chapter we have been mainly concerned with methods for testing hypotheses. Tests we described, as well as all other such tests, are based upon certain assumptions. For example, most applications of operating characteristics involve an assumption of normality, and most nonparametric techniques are based upon the theory of equally-likely permutations. The assumptions behind a particular statistical technique constitute a *model*. Whenever a psychologist makes a set of measurements on two groups of people he implicitly employs a rather general model which says that each set of measurements comes from a population and each population can be represented by some kind of a distribution function. Then he usually hypothesizes that the two populations have different means or different variances or different something-else. But when he uses a particular statistical technique he implicitly or explicitly makes further assumptions about the two distributions and so is using a more specific model.

Testing of hypotheses is one of the main branches of statistics. Another is the estimation of parameters. Whenever a social psychologist draws a sample and makes inferences about the population from which it was drawn, he is estimating a parameter, e.g., the population mean or variance. Here again there is an underlying model; one necessarily makes some assumptions about the parent population, for otherwise the parameter estimated has little meaning or interest. To be sure, the model may be a very simple one, but it is still a model. There are many problems in estimation, however, where the model is not so simple, e.g., the analysis of time series.

An outstanding feature of statistical models —whether for testing hypotheses or estimating parameters—is that they are selected to suit each experimental problem as it arises. The statistician has provided the psychologist with a rather large collection of models to choose from. The experimenter looks at his data, or preferably a projected experiment, and tries to find a suitable model. Moreover, he does not have the slightest hesitation to use an entirely different model for the next experiment. In this sense, the use of statistical models is *data-oriented*. We next discuss a different class of models which are *theory-oriented*.

The term *behavioral model* is now often used to refer to a somewhat different type of application of mathematics to psychology. The behavioral model is not so much designed to help in the analysis of a particular set of data as in the construction of theory. Hull's attempt to formalize rote learning is an example. When one attempts to "explain" a large class of phenomena — to tie together a broad range of behavior—he is building a behavioral model. The use of mathematics in such model construction is not at all essential, but there is a growing interest in these attempts. The distinction beween behavioral and statistical models neither is nor needs to be a very clear one. A behavioral model, if it is a good one, should suggest ways of handling data, and a statistical model often provides stimulation to theoreticians. The difference is mainly in orientation, not use. While the statistical model is chosen to analyze a particular set of data, the behavioral model should lead to the acquiring of new kinds of data or at least to the selection of data which is relevant to it; the "testing" of the behavioral model is of primary interest.

It is interesting to note an important difference in the role of significance tests when such models are tested. Instead of searching for significant differences and being pleased by finding them, the model builder is hoping that the null hypothesis can be accepted, and that his model does fit the data.

In closing we would like to remind the reader that most statistical and mathematical model problems have yet to be investigated, and that his inquiry to the nearest statistician or mathematician for help will not always bring a ready answer. Indeed, social psychology is a prominent source of research problems for the mathematical statistician and perhaps is becoming such a source for the applied mathematician.

REFERENCES

ADLER, F. Yates' correction and the statisticians. *J. Am. statist. Ass.*, 1951, *46*, 490-501.

BATTIN, I. L. On the problem of multiple matching. *Ann. Math. Statist.*, 1942, *13*, 294-305.

COCHRAN, W. G. Analysis of variance for percentages based on unequal numbers. *J. Am. statist. Ass.*, 1943, *38*, 287-301.

COCHRAN, W. G. The χ^2 test of goodness of fit. *Ann. Math. Statist.*, 1952, *23*, 315-346.

DIXON, W. J., & MOOD, A. M. The statistical sign test. *J. Am. statist. Ass.*, 1946, *41*, 557-566.

DUNCAN, D. B. A significance test for differences between ranked treatments in an analysis of variance. *Va. J. Sci.*, 1951, *2*, 171-189.

DUNCAN, D. B. On the properties of the multiple comparisons test. *Va. J. Sci.*, 1952, *3*, 49-67.

EDWARDS, A. L. *Statistical analysis.* New York: Rinehart, 1946.

EDWARDS, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.

EICHLER, R. M. Some comments on the controlling of differences in responses on the Rorschach test. *Psychol. Bull.*, 1951, *48*, 257-259.

FERRIS, C. D., GRUBBS, F. E., & WEAVER, C. L. Operating characteristics for the common statistical tests of significance. *Ann. Math. Statist.*, 1946, *17*, 178-197.

FISHER, R. A. *Statistical methods for research workers.* (7th Ed.) London: Oliver & Boyd, 1938.

FREEMAN, M. F., & TUKEY, J. W. Transformations related to the angular and the square root. *Ann. Math. Statist.*, 1950, *21*, 607-611.

FRIEDMAN, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. statist. Ass.*, 1937, *32*, 675-701.

GORDON, M. H., LOVELAND, E. H., & CURETON, E. E. An extended table of chi-square for two degrees of freedom, for use in combining probabilities from independent samples. *Psychometrika*, 1952, *17*, 311-316.

GUILFORD, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.

Hamilton, C. Questions and Answers: Comment on Question 8, Variance of number of matches. *Am. Statistician,* 1948, *2,* 23.

Kendall, M. G. *The advanced theory of statistics.* Vol. I. London: J. B. Lippincott, 1943.

Kendall, M. G. *Rank correlation methods.* London: Charles Griffin, 1948.

Kruskal, W. H., & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Am. statist. Ass.,* 1952, *47,* 583-621.

Link, R. F., & Wallace, D. L. Some short cuts to allowances. Unpublished manuscript, Princeton University.

Lord, E. The use of range in place of standard deviation in the *t*-test. *Biometrika,* 1947, *34,* 41-67.

McNemar, Q., & Terman, L. M. Sex differences in variational tendency. *Genet. Psychol. Monogr.,* 1936, *18,* No. 1.

MacStewart, W. A note on the power of the sign test. *Ann. Math. Statist.,* 1941, *12,* 236-238.

Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.,* 1947, *18,* 50-60.

Mantel, N. Means for small samples. *Am. Statistician,* 1951, *5,* 26-27.

Mood, A. M. *Introduction to the theory of statistics.* New York: McGraw-Hill, 1950.

Moses, L. E. Non-parametric statistics for psychological research. *Psychol. Bull.,* 1952, *49,* 122-143.

Moshman, J. Testing a straggler mean in a two-way classification using the range. *Ann. Math. Statist.,* 1952, *23,* 126-132.

Pearson, K. (Ed.) *Tables for Statisticians and Biometricians. Part II.* London: Biometrika Publications, 1931, 78.

Scheffé, H. A method for judging all contrasts in the analysis of variance. *Biometrika,* 1953, *40,* 87-104.

Snedecor, G. W. *Statistical methods.* Ames, Iowa: Iowa State College Press, 1946.

Stouffer, S. A. *et al. The American soldier. Vol. I. Adjustment during army life.* Princeton: Princeton University Press, 1949.

Swed, Frieda S., & Eisenhart, C. Tables for testing randomness of grouping in a sequence of alternatives. *Ann. Math. Statist.,* 1943, *14,* 66-87.

Tocher, K. D. Extension of the Neyman-Pearson theory of tests to discontinuous variates. *Biometrika,* 1950, *37,* 130-144.

Tukey, J. W. Comparing individual means in the analysis of variance. *Biometrics,* 1949, *5,* 99-114.

Tukey, J. W. Quick and dirty methods in statistics. Part II. Simple analyses for standard designs. *Proc. 5th Annual Convention, Amer. Soc. for Quality Control,* 1951, 189-197.

Wallis, W. A. Compounding probabilities from independent significance tests. *Econometrica,* 1942, *10,* 229-248.

Wallis, W. A. Statistics of the Kinsey report. *J. Amer. statist. Ass.,* 1949, *44,* 463-484.

Whitney, D. R. A bivariate extension of the U statistic. *Ann. Math. statist.,* 1951, *22,* 274-282.

Wilcoxon, F. Individual comparisons by ranking methods. *Biometrics Bull.,* 1945, *1,* 80-82.

Wilks, S. S. *Mathematical statistics.* Princeton: Princeton University Press, 1943.

Tables of the binomial probability distribution. Washington: Department of Commerce, National Bureau of Standards, Applied Mathematics Series 6, 1950.

CHAPTER 9

Attitude Measurement

BERT F. GREEN
Massachusetts Institute of Technology

In this chapter we shall discuss methods for developing meaningful quantitative measures of the attitudes of individuals. Our purpose is to present the logical basis of attitude measurement in general, and to give a detailed account of the prominent attitude-scaling techniques.

Since McNemar (1946) wrote his comprehensive critical review of opinion-attitude methodology, some major methodological advances have been reported by Stouffer, Guttman, Suchman, and Lazarsfeld (Stouffer, *et al.*, 1950a) in connection with attitude research in the U. S. Army. This work has greatly influenced the formulations in this chapter. Another important influence, which the author gratefully acknowledges, is W. S. Torgerson's monograph on the theory and methods of scaling, being prepared for the committee on scaling of the Social Science Research Council. Manuscript for portions of this monograph was kindly made available by Dr. Torgerson. Specific citations will not be made in this chapter.

THE CONCEPT OF ATTITUDE

To be able to understand how attitudes can be measured, we first need to examine the concept of attitude. Like many psychological variables, attitude is a hypothetical or latent variable, rather than an immediately observable variable. The concept of attitude does not refer to any one specific act or response of an individual, but is an abstraction from a large number of related acts or responses. For example, when we state that individual *A* has a less favorable attitude toward labor organizations than individual *B*, we mean that *A*'s many different statements and actions concerning labor organizations are consistently less favorable to labor than are *B's* comparable words and deeds. We are justified in using a comprehensive concept like attitude when the many related responses are consistent. That is, if people who disapprove of the closed shop are also likely to want to outlaw strikes, and to oppose minimum wage laws, then it seems reasonable to speak of an antilabor attitude.

In general terms, a latent variable is used to describe the consistency or covariation of a number of different responses to stimuli of the same general class. The variable is viewed as mediating the stimuli and the responses. The responses are said to covary because they are all mediated by the same hypothetical variable. The latent variable is useful because it unifies a set of data, namely, the observed responses. Hypothetical variables of this kind have been called *traits, intervening variables* (MacCorquodale and Meehl, 1948), *latent variables* (Lazarsfeld, 1950a), *genotypes* (Coombs, 1952), and *factors* (Thurstone, 1947). The observable data have been called *manifest variables, phenotypes,* and *indicants* (Stevens, 1951).

Many definitions of social attitudes have been proposed. These definitions seem to have a common theme that suggests the manner in which attitude may be viewed as a latent variable. This theme is the concept of attitude as a consistency among responses to a specified set of stimuli, or social objects.

Consider a few examples. Krech and Crutchfield (1948, p. 152) view attitude as ". . . an enduring organization of motivational, emotional, perceptual, and cognitive processes with respect to some aspect of the individual's world." An enduring organization of psychological processes seems to imply a consistency of response patterns. For Doob (1948) an attitude refers to an implicit response that is both anticipatory and mediating in reference to patterns of overt responses, that is evoked by a variety of stimulus patterns, and that is considered socially significant in the individual's society. Overt responses are consistent in that they are mediated by the implicit response. Allport (1935, p. 810), after reviewing many early definitions of attitude, concluded that "an attitude is a mental and neural state of readiness exerting a directive influence upon the individual's response to all objects and situations with which it is related." Fuson (1942) defines an attitude as the probability of occurrence of a defined behavior in a defined situation. Campbell (1950) suggests that a social attitude is evidenced by consistency in response to social situations.

It is apparent from these examples that the concept of attitude implies a consistency or predictability of responses. An attitude governs, or mediates, or predicts, or is evidenced by a variety of responses to some specified set of social objects or situations. Campbell (1950, p. 31) has summarized this view neatly in presenting an operational definition of attitude: *"An individual's social attitude is an* [*enduring*] *syndrome of response consistency with regard to* [*a set of*] *social objects."*

This definition does not divest attitudes of their affective and cognitive properties, which may be properties of, or correlates of, the responses that comprise the attitude. However, attention is focused on the characteristic of attitude that is basic to all attitude measurement: response covariation. In each measurement method, covariation among responses is related to the variation of an underlying variable. The latent attitude is defined by the correlations among responses.

The set of social objects that forms the reference class of an attitude distinguishes attitude from other psychological variables such as habit, temperament, drive, or intelligence. It is of secondary importance whether we call the variable an attitude, or a trait, or a habit. The operational definition will always be in terms of the referent class of stimuli.

The content of an attitude is determined by the responses which constitute it. The set of behaviors comprising an attitude will be called an *attitude universe.* [Guttman (1950a) developed the idea of a universe in connection with his scaling method; his approach seems applicable to the general measurement problem.] The elements of this universe are manifest variables—the responses to specific situations. A sample of these elements is used to measure the attitude; from this sample of behavior an inference can be made about the entire universe. For example, if a sample of elements is homogeneous, it is reasonable to infer that the universe also is homogeneous. Furthermore, the scores of individuals on this scale may be taken as representative of their behavior as described by the universe of elements. On the other hand, homogeneous subsets of elements may be found within the original universe, implying that several dimensions are represented in the original attitude universe.

Conceptually, then, attitude measurement involves sampling a behavior universe, and measuring the universe by means of the sample. This means that the sample of elements should be representative. If only a small subset of the total attitude universe is sampled, then inference beyond this subset is not legitimate. Of course, in practice we seldom have the universe, but only the sample. We may characterize the elements of the hypothetical universe by means of the sample; however, it is very important not to overgeneralize. From a sample of verbal responses to questions about opinions, one should not make inferences about behavior other than verbal responses to similar verbal questions. It may be that responses to these verbal questions are correlated with responses in nonverbal situations, but this must be determined experimentally.

The concepts of attitude universe and subuniverse are certainly not precise, but they are helpful in formulating the measurement problem, and will be particularly useful in discussing validity.

As we have defined it, an attitude is a latent variable, since its meaning is derived from the covariation of responses in some attitude universe. To obtain a more precise definition of an attitude, we need a mathematical model that relates the responses, or observed variables, to the latent variable. Each psychological scaling method either states or implies such a model. The problem of measuring an attitude is one of selecting a scaling model by which the response data can be related to the attitude variable.

PROPERTIES OF SCALES

Different scaling methods have different properties. To select a scaling method wisely, it is important to understand the various properties of scales of measurement.

The quantification of psychological attributes has been the subject of much debate. Some scientists, including Ferguson, Myers, Bartlett, *et al.* (1940), Bergmann and Spence (1944), Johnson (1936), have argued that much psychological measurement is not true measurement, since it does not conform to the methods and definitions used in physical measurement. They point out that most physical scales have the property of additivity, and that this property may be demonstrated experimentally. An operational definition of the sum of two lengths may be achieved by laying two sticks end to end; in the case of weight, two weights and a balance are needed. By what psychological experiment, it is asked, may two pitches be added, or two attitudes?

One answer has been given by Gulliksen (1946). He showed that a logical extension of the concept of additivity will allow scales devised by the method of paired comparisons to be classed as additive. In this case the difference between two scale values is defined operationally. The experiment consists in determining the proportion of trials on which stimulus i is judged to be greater than stimulus j with respect to the attribute in question. A transformation based on the normal distribution relates this observed proportion to the difference between the scale values of the two stimuli. If the differences between the scale values of all pairs of stimuli are mutually consistent, e.g., $(S_i - S_j) + (S_j - S_k) = (S_i - S_k)$, then the additive property of the scale is established.

The argument about the properties of a scale of measurement has led to a broader conception of the nature of measurement. Stevens (1951), Lorge (1951), and others have pointed out that in its broadest sense, *measurement is the assignment of numerals to objects, events, or persons, according to rules.* With the rules properly defined, the numerals may be used to represent relations among the objects, and to compare these objects with other sets of objects similarly measured. The result of a measurement is a *scale.* A scale comprises the set of numerals given to the objects by using a certain rule of assignment. The objects are then said to be scaled.

Objects may be measured according to many different rules, each of which leads to a scale with special properties. Stevens (1951) distinguishes among *nominal, ordinal, interval,* and *ratio* scales. In a nominal scale, objects are placed in several mutually exclusive categories, but there is no necessary relation among the categories. The number assigned to the objects in a category serves as a label, but has no quantitative meaning. Any other set of labels would be adequate if each original label is replaced by a different label in the new set (i.e., a one-to-one transformation). When objects can be arranged in a rank order with respect to some variable, they are said to form an ordinal scale. Numbers may be assigned to the objects if the numbers are in the same rank order as the objects. (Any monotonic transformation is permissible.) We might obtain an ordinal scale of the hardness of the several solid substances by determining which substances scratch which other substances. When the intervals between objects on a scale can be measured, then an interval scale exists. An interval scale of hardness would require some experimental method to determine *how much* harder substance A was than substance B. We could then assign a number to each substance so that the difference between the two values indicated the extent of the difference in hardness. Another set of numbers could be substituted for the original set only if there was a linear relationship ($y = ax + b$) between the two sets of numbers. A ratio scale results when there is some way of showing how many *times* greater one object is than another. Numbers are assigned to the objects so that the numbers have the same ratio as the experimentally determined ratios of the objects. A ratio scale implies a fixed zero point, so that the only admissible transformation is multiplication by a constant ($y = ax$). Weight forms a ratio scale because we can show that four pounds is four times as heavy as one pound, and similarly for all pairs of weights.

This classification is not complete, although it includes the main properties of scales. Coombs (1950, 1952) has suggested several scales that fall between nominal and ordinal scales. For example, we could have an ordered set of categories, with objects within a category being unordered. A partially ordered set of objects would result if some objects could not be ordered with respect to some other objects. Coombs has also suggested an "ordered metric," in which the objects are ordered, and the intervals between objects on the scale are also ordered.

Ordinal, interval, and ratio scales are defined with respect to a single dimension along which the objects are placed. These definitions can be extended to include multidimensional scales. It is an experimental question whether the relations among the objects are consistent with a one-dimensional continuum. For example, we might find an inconsistent set of differences, such as $A - B = 4$, $B - C = 3$, $A - C = 5$. Here, if the result is not to be called "error," the quantities must be referred to a two-dimensional scale. A, B, and C could be represented as points on a plane, with the obtained differences taken as the distances between the points, whereupon A, B, and C would be the vertices of a right triangle. For example, in a scale of colors, A and B might differ only in saturation, and B and C might differ only in hue, while A and C might differ in both saturation and hue. In this case, A and C would appear to be more dissimilar than A and B, or B and C. Although multidimensional scaling will not be treated here (see Torgerson, 1952, and Bennett, 1951), the general concept is basic to the various definitions of unidimensionality that will be met later.

DEFINITIONS

The terms *measurement, scale, attitude,* and *attitude universe* have been defined above. In this section we give definitions for other terms which we shall use in special senses.

Attitude measurement consists in assessing an individual's responses to a set of social objects or situations. This is done by observing a sample of behavior from an attitude universe. Each behavioral element in the attitude universe is the response to a particular situation or object. The particular situation or object that evokes the response, together with a specified set of response categories, is called an *item*. An item is a manifest variable. The statement, "The government should own the coal mines," together with two response categories, "agree" and "disagree," forms an item. The same statement, with five response categories, "strongly agree," "agree," "neutral," "disagree," "strongly disagree," is a different item. In the statistical analysis, the response categories to the latter may be combined to produce a dichotomous variable, but from the respondent's point of view, the two items require slightly different kinds of responses.

A number assigned to an individual's item response is an *item score*. In the first example, "agree" might be scored 1, and "disagree," 0. The proportion of people who give a certain response to an item determines the *popularity* of that response. The popularities of the various responses to an item are the *item marginals.* For dichotomous items, it will be convenient to specify arbitrarily one of the response alternatives as the positive response. It will be sufficient to consider only the popularity of the positive response, and thus a single marginal for each item. The popularity of a positive response to item i will be symbolized by p_i. We shall also use p_{ij} for the proportion of people responding positively to both items i and j, p_{ijk} for the proportion responding positively to all three items i, j, and k, etc.

Any collection of items will be called a *questionnaire.* Other terms, such as test, inventory, and scale, have been used in the literature. The term *questionnaire* is free from some of the connotations of the other terms. A questionnaire may contain items from different attitude universes, although usually only a single universe is represented. The distinction will be clear from the context.

An individual whose attitude is being measured will be called a *respondent.* A number, derived from his item responses, that represents his position on the latent variable is called a *scale score*, or simply a *score*. The scores form the *attitude scale*. The scale implies a system for obtaining a score from an individual's item responses. The term *scale* will occasionally be used to refer to this scoring system.

THE EVALUATION OF SCALES

To evaluate a scale, statistical evidence is needed concerning the properties of the scale. The *reliability* of the scale is of primary importance. A statement should also be provided about the *attitude universe* represented by the scale — that is, what items might have been used. Evidence of correlational *validity* is of paramount interest, since the utility of the scale will depend on its relationship with other variables and its ability to predict other behavior.

Facts such as these are necessary to specify what is meant by the label that is attached to the scale, and to ensure against false interpretations of that label. Putting a name on a scale creates a serious social obligation that cannot be fulfilled without experimental evidence concerning the properties of the scale.

Reliability. The reliability of the scale is an index of the extent to which repeated measurements yield similar results. We shall not discuss the reliability of group statistics such as means and proportions, where discrepancies are due to sampling different individuals. Rather, our primary concern is with the reliability of an individual's score. If chance fluctuations cause relatively large shifts in an individual's score, then any particular determination of the score is practically meaningless. We would not expect such scores to correlate highly with other variables. Since attitude scales are developed as research tools for studying the relation of attitudes to other variables, high reliability is an indispensable property of a scale.

Reliability has two aspects, which Cronbach (1951) has called *stability* and *equivalence.* These are new names for test-retest reliability and parallel-form reliability, respectively.

If the same questionnaire is administered to the same group of respondents on two separate occasions, there will be some shifts in the scores of individuals. The correlation between these scores is the test-retest reliability or the coefficient of stability. Memory and familiarity with the scale will tend to increase stability; extending the time interval between administrations will tend to decrease stability.

Low stabilities of attitude scales have been defended on the grounds that attitudes are dynamic and should be expected to change.

This may be true, but it is difficult to see how an attitude scale can be a useful predictive instrument if the scores are not stable. Furthermore, there is a considerable difference between consistent shifts in response syndromes and sporadic changes in specific response tendencies. In studies of attitude change we are concerned with consistent shifts of response syndromes. In this case, low stability of the score is expected. But it is important to know the reliability of the *change*. This can be discovered by means of equivalent scales.

If our purpose is to measure an attitude universe by means of a sample of items, then we must determine what score differences could be expected if a different sample of items were chosen. Scales based on equivalent samples of items from the same universe should yield very similar scores. The correlation between such "parallel" scales is called the coefficient of equivalence, or the parallel-form reliability. It measures the extent to which the scale is specific to the particular items used. We are not justified in generalizing beyond the specific items in the scale unless equivalence is demonstrated. The concept of an attitude universe implies that there are many items which might have been used in the questionnaire. If, in fact, an equivalent set of items cannot be produced, the notion of an attitude universe becomes meaningless. Of course, in some situations we are interested in a scale for a specific set of items, but more often we would like to generalize to a larger class of possible items.

Two scales derived from different samples of items from the same universe should be highly correlated. However, we would also like the two scales to have similar scale-score distributions, similar patterns of inter-item correlations, and similar reliabilities. Such equivalent, or parallel, scales are best constructed from matched or paired samples of items. If possible, items with similar response distributions should be paired, i.e., items with equal marginals. Paired items should also have a similar degree of relationship with the total scale. Parallel scales constructed in this way, if they have high coefficients of equivalence, can be used interchangeably.

There are several methods for estimating the coefficient of equivalence from a single set of items. These methods depend on the strength of the interrelationships among the items, that is, on the *homogeneity* of the scale. The most familiar of these methods is the split-half technique. In this technique, the items are divided into two subsets, either matched or random (odd-even), and a scale is developed for each subset. The correlation between these scales may be used with the Spearman-Brown formula to estimate the equivalence of the original scale. (Technically, the Spearman-Brown formula is applicable only when the scale score is a weighted or unweighted sum of the item scores.) Guttman (1945, 1946) has shown that his coefficient of reproducibility is in some cases a lower bound of equivalence. The Kuder-Richardson formulas, especially formula (20) which Cronbach (1951) has called coefficient α, have been widely used as estimates of equivalence. Loevinger's (1947) coefficient of homogeneity is a direct index of the homogeneity of the items; it does not estimate the reliability of the scale. A discussion of some of these methods, as well as a thorough treatment of reliability, may be found in Gulliksen (1950b).

Unidimensionality and homogeneity. A unidimensional scale measures a single variable. In general, the property of unidimensionality means that people with the same score have about the same attitude. If the scale is actually a reflection of more than one attitude, so that it acts as the combination of two or more scales, then the same score can represent very different attitude patterns. For example, a scale of ethnocentrism composed of five items concerning attitudes toward Negroes, five items concerning attitudes toward Italians, and five items concerning attitudes towards the British would probably not be homogeneous. A median score might indicate median attitudes toward all three groups, or a strong anti-Negro attitude combined with a median attitude toward Italians and a strong pro-British attitude, or any other similar combination. Of course, if the three attitudes are highly correlated, then such a scale could be considered homogeneous and unidimensional, since ambiguous interpretations of scores would be improbable.

Operationally, unidimensionality and homogeneity have essentially the same meaning. Both are defined in terms of the degree of covariation among the items. In general, if the items on a scale are highly interdependent, then the items are homogeneous and the scale is unidimensional. More specific considerations must be discussed in connection with particular scaling methods, since each method contains its own definition of unidimensionality.

A homogeneous scale is, in general, a reliable scale. In the previous section we noted that several methods of estimating the coefficient of equivalence are based on homogeneity. In these methods it is assumed that the degree of relationship or covariation among the items in a scale is representative of the degree of covariation to be expected between these items and an equivalent set of items. However, it is possible for two highly correlated parallel scales not to be homogeneous. The scale described in the preceding paragraph could be highly reliable yet not homogeneous. Thus, if equivalence is measured by correlating two equivalent scales, additional evidence of homogeneity is needed.

Validity. If a scale is reliable and homogeneous, it measures some variable. The problem of validity is to discover what it measures. If the scale lives up to our definition of attitude, then it measures an attitude universe. The first problem in validation is to delimit and define this attitude universe.

It is helpful to distinguish at least three kinds of attitude universes. Most attitude questionnaires consist of items in which the respondent is asked to give some verbal response. This verbal response is quite often a statement of the respondent's opinion, or some other form of self-description. These attitude universes may be called *elicited verbal attitudes.* The opinions expressed by individuals in normal conversation with friends may be classed as elements of *spontaneous verbal attitudes.* A third class of behavior is verbal or nonverbal behavior directed toward an object in the referent class. We may call the corresponding universe an *action attitude.* For example, a restaurant manager might tell a survey interviewer that he would not serve Negroes, he might assure his friends that he won't serve Negroes, and he might refuse service to a Negro in his restaurant. It might also happen that the three responses are not so consistent.

Responses in any of these areas may be scaled. Although most of our attitude scales measure elicited verbal attitudes, other scales are equally possible. With clever observational techniques, scales could be developed for spontaneous verbal attitudes and for action attitudes.

These classes of attitude are not exhaustive. The indirect techniques lead to further attitude universes. Also, subclasses can be made within the three stated classes. Action attitudes are particularly diverse, but little measurement has been done in that area, so that breakdowns are as yet unnecessary.

A description of the attitude universe is necessary in order to determine whether the scale measures this universe. Essentially, this means characterizing the items which might have been used in the questionnaire. Many investigators would say that, as a representation of an attitude universe, the scale is valid by definition, or has face validity. Gulliksen (1950a) has used the term *intrinsic* validity for this situation, and has suggested some experimental methods for determining whether a scale is intrinsically valid. One method is to study the judgments of a group of experts. A large set of items could be assembled which the experimenter believes to be in the same attitude universe as the items on the scale. Experts could rate these items as well as the items on the scale for relevance to the attitude being considered. Alternatively, several experts could each submit a questionnaire that is believed to be a measure of the attitude universe. These questionnaires could be administered and their intercorrelations investigated by factor analysis to see if a single common factor is exhibited by the several scales. Such techniques promise to give at least a partial solution to the plaguing problem of attitude scale validity.

The second major method of determining what a scale measures is to find the variable with which it correlates. By definition, an attitude scale for one universe cannot *measure* a second universe. However, if it correlates highly with the second universe, then it may legitimately be used as a substitute for such a measure. In particular, an elicited verbal attitude scale can never be a measure of an action attitude, but it may be taken as an indicant of the action attitude if a reasonably high correlation exists between the two. Since this is often the aim of elicited verbal attitude scales, correlation with an action attitude is called its validity for the action attitude or just its validity.

Unfortunately, it is difficult to determine whether such a correlation exists. The validity of elicited verbal attitudes has often been appraised by selecting two groups of people who are judged to have different action attitudes, and then demonstrating that their elicited verbal attitudes differ in the expected direction. A major difficulty of this *known group* validation is that the magnitude of the relationship cannot be easily evaluated. Furthermore, the groups are seldom very well "known." The ideal method of determining correlational validity would be to develop an action attitude scale, which might consist of a set of standardized social situations, or a behavior checklist which would require extensive observation of each individual.

Many investigations have found that specific acts or action attitudes often cannot be predicted very accurately from elicited verbal attitudes. The classic demonstration of the lack of correspondence of verbal and nonverbal behavior is that of LaPiere (1934). He accompanied a Chinese couple on some of their travels, and later queried the managers of the restaurants, hotels, and auto camps that they had visited. Over 90% said they would not accept Chinese guests, whereas all of them had done so.

Emphasis on the validity of verbal attitude scales for predicting action attitudes often obscures the fundamental issue in attitude research. We are interested in the relation of attitudes to other variables, such as socioeconomic status, education, exposure to propaganda, or other attitudes. If we wish to study the relation of elicited verbal attitudes to such variables, then we need a scale that measures elicited verbal attitude. If we want to investigate the correlates of spontaneous verbal attitudes, or action attitudes, then we need scales that measure these attitudes directly. It is only when we attempt to use a scale based on elicited verbal attitude as an indicant of action attitude that this question of validity is paramount.

ITEMS AND QUESTIONNAIRES

Items are the elements from which attitude scales are constructed. The usefulness of a scale, as well as the definition of the attitude that it measures, depends ultimately on the items of which it is composed. Ingenuity and insight are at a premium in designing items that will yield valid and useful scales.

There are many types and forms of items that may be used to assess an attitude. Item *type* refers to the content of the item, while item *form* refers to the manner in which the item is related to the latent continuum. The scaling methods are largely indifferent to the type of item used, although certain requirements are made concerning the form of the item. The scaling methods usually include procedures for detecting and eliminating items which do not meet these formal requirements.

ITEM TYPES

A great many types of items have been used in attitude research. Although the scaling methods are usually used with opinion statements, this is not necessarily the case. Almost any item can be used if it has a specified set of possible responses. Although we shall not attempt a survey of all possible item types, we shall give a few examples to illustrate the range of item types that can be encompassed by the scaling techniques.

The most widely used item type is the opinion statement. The respondent may be asked whether he agrees with the statement, or he may be required to indicate his opinion on a five-point scale from "strongly agree" to "strongly disagree." Occasionally the respondent is asked which of several explicitly presented opinions on the issue is most nearly like his own opinion. Open-ended questions may be used. Here the respondent expresses his own opinion, and the response is later placed in one of a fixed number of categories. Thurstone and Chave (1929) suggested that opinion statements could be obtained for objective questionnaires by collecting the responses to open-ended opinion questions.

The responses to opinion statements are notoriously sensitive to changes in the wording of the statement. This is a more serious problem in single-question opinion polling than in attitude scaling, where the use of several statements can partially offset the effect. However, this does not constitute a license for sloppy wording in attitude items. In a readable book, *The Art of Asking Questions,* Payne (1951) discusses the many pitfalls to be avoided in preparing opinion statements or questions. Payne (1950, 1951) also discusses the problem of meaning. People are often willing to state opinions about issues that are meaningless to them. It is important that the opinion statement be relevant to the respondent as well as to the attitude. This issue is treated in greater detail in Chapter 12.

The use of opinion statements is tantamount to asking the respondent for a structured description of his verbal attitude. In contrast, Rosander (1937) asked for a description of the respondent's action attitude. He posed several hypothetical situations and described a particular action; the respondent indicated whether he would have taken that action. One item from a scale of attitudes toward Negroes is: "You are reading in a public library. A Negro enters and sits down beside you. You leave the library at once." Pace (1939) used similar items for general social-economic-political attitudes. F. H. Allport and Hanchett (1940) used this type of item to study war-producing behavior. Validities found with these action-oriented items indicate that they are very promising.

A similar idea was used by Horowitz and Horowitz (1938) in studying the development of attitude toward the Negro in young children. In their "Faces" test the child was shown a set of pictures of white and colored children and was asked to point out the one he wanted to play with, to eat with, etc. The child could also be asked to rank the faces in order of preference. Mussen (1950) reports a validity of .52 for this test, using a criterion of sociometric choice at an interracial summer camp.

The methods for the selection of OSS personnel developed during the war by the OSS assessment staff (1948) could be adapted for use in action attitude research. Essentially the technique would involve establishing a series of standardized social situations relevant to the attitude in question, and studying the behavior of the individual in these situations. One could presumably establish an objective scale based on the responses in a number of these standard social situations. (Kelley and Fiske, 1951, used these methods to study the selection of students in clinical psychology.)

Action attitudes can be measured by behavior rating scales. The items could be the determination of presence or absence of a certain action in a particular social situation. Such scales would be similar to a scale based on standard social situations. In this case, however, the scales would depend on raters who knew the individual well. The rater could be the individual himself or some close acquaintance. Great care must be taken to make the rater's job objective, so that reliable results can be expected.

When the purpose of the attitude questionnaire is disguised, the measurement has been called *indirect.* Campbell (1950) and Weschler and Bernberg (1950) have provided comprehensive reviews of these indirect methods. At least two major categories of indirect items can be distinguished: items of the projective type and items based on the selective distortion of psychological processes.

Several investigators have used items similar to those on the Thematic Apperception Test. Proshansky (1943) used ambiguous pictures of labor situations to measure attitude toward labor. The measurement was a rating made on the basis of the respondent's written interpretations of the pictures. Fromme (1941) used political cartoons with four alternative captions. The respondent selected the best caption for each cartoon, and gave reasons for his selections. Although Fromme's analysis was qualitative, an objective scale could be developed easily. Sanford (1950) studied worry by using items similar to those on the Rosenzweig Picture Frustration Test. For example, in a cartoon depicting two men, one is asking the other what he is worried about. The respondent is required to supply the second man's response.

Several indirect methods are based on the premise that attitudes interfere with psychological processes such as learning, perception, and remembering. The selective distortion of these processes is a prominent phenomenon in social psychology. Consistency among such distortions is closer to some traditional conceptions of attitude than is consistency among opinions.

Several investigators have used items based on the distortion of actual or assumed facts. Smith (1947) used statements labeled as facts; the respondent indicated whether he believed the statement to be true. Newcomb (1946) found that errors on an information test were related to the respondent's attitude. In Hammond's (1948) "error choice" technique, a factual question is asked, with two response alternatives available, each equally far from the truth, but in different directions. For example, "The number of man-days lost because of strikes from January to June was (1) 34.5 million, (2) 98.6 million." Bernberg (1951) devised a scale of industrial morale based on this method of "error choice."

The selective distortion of logical reasoning has been used by Thistlethwaite (1950) as a basis for attitude items. He paired neutral and "loaded" items of the same (valid or invalid) syllogism. The subject is asked to indicate for each item whether the conclusion is justified by the premise. An example of a "loaded" (invalid) syllogism is: "Given: If production is important, then peaceful industrial relations are desirable. If production is important, then it is a mistake to have Negroes for foremen and leaders over whites. Therefore: If peaceful industrial relations are desirable, then it is a mistake to have Negroes for foremen and leaders over whites."

Indirect attitude items may also be based on distortion of perception and memory. Horowitz and Horowitz (1938) invented an *Aussage Test,* which might also be titled "The 'Have you stopped beating your wife' Test." A complicated picture is presented for two or three seconds, after which the respondent is asked misleading questions such as "What is the colored man in the corner doing?" Stereotyped responses are often obtained concerning the nonexistent colored man.

That the relation between these indirect items and scales based on opinion statements may be complex is shown by Bray's (1950) study of prejudice. He used Sherif's (1937) autokinetic situation with confederates who were Negroes

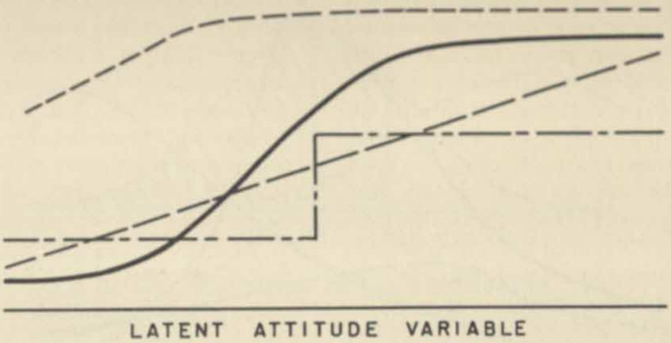


FIG. 1. Examples of monotone item operating characteristics. The ordinate is the expected item score (the probability of a positive response, for dichotomous items).

or who were introduced to the subjects as Jews. The amount of conformity of subjects' reports to confederates' reports could not be directly predicted by scores on attitude scales composed of opinion statements. Rather, the interaction of the attitude scores and personality variables such as irritability, social ascendancy, and self-confidence were of major importance in determining the conformity of reports.

ITEM FORM

Each scaling method makes certain requirements about the relationship of the manifest item to the latent attitude variable. This relationship can best be described in terms of the average item score for a given scale score, that is to say, the regression of item score on the latent variable. This curve is called the *item operating characteristic.* (In the preceding chapter the term *operating characteristic* is applied to the curve indicating the probability of a type II error in a statistical test. In general, the term refers to a quantitative statement of the effects to be expected when using some procedure or object, such as an attitude item, a statistical test, or a vacuum tube.) For a dichotomous item, if the positive response is scored 1 and the negative response 0, then the average item score is equivalent to the probability of a positive response. In this case the item operating characteristic gives the conditional probability of a positive response for any fixed value of the latent variable. Lazarsfeld (1950a) uses the term *trace line* in referring to the item operating characteristic.

The so-called psychometric function in psychophysics is very similar to the item operating characteristic, in the dichotomous case. Mosier (1940b) has pointed out that the psychometric function can be translated into the item operating characteristic by interchanging the roles of persons and items, and vice versa. The psychometric function indicates the probability that a person will perceive a stimulus, as a function of the stimulus intensity, whereas the item operating characteristic indicates the probability that a person will agree with, or pass, an item, as a function of the person's attitude.

There are two major forms of items: *monotone* and *nonmonotone.* We are using Coombs' (1952) terminology; Stouffer (1950b) and Mosteller (1949) use the terms *cumulative* and *noncumulative.* Items in the two classes may be distinguished by the general shape of their operating characteristics. Monotone items have monotonic characteristics. That is, as the value of the latent variable increases, the operating characteristic increases, or may remain constant, but never decreases. Some examples are shown in Fig. 1.

Examples of monotone items, for a scale of height, are: "Are you taller than a Chevrolet?" or "Are you taller than 6′0″?" Note that if an item is scored in the reverse direction, so that its operating characteristic decreases or remains constant, but never increases as the latent variable increases, it is still considered a monotone item.

The operating characteristics of nonmonotone items, on the other hand, first increase, attain a maximum, and then decrease, as the latent variable increases. Examples of such characteristics are shown in Fig. 2.

Examples of nonmonotone items are: "Are you about as tall as a Chevrolet?" or "Are you between 5′9″ and 6′3″ tall?" We have considered only nonmonotone items with single maxima. Occasionally an item may appear to have two or more maxima. The "double-barreled"

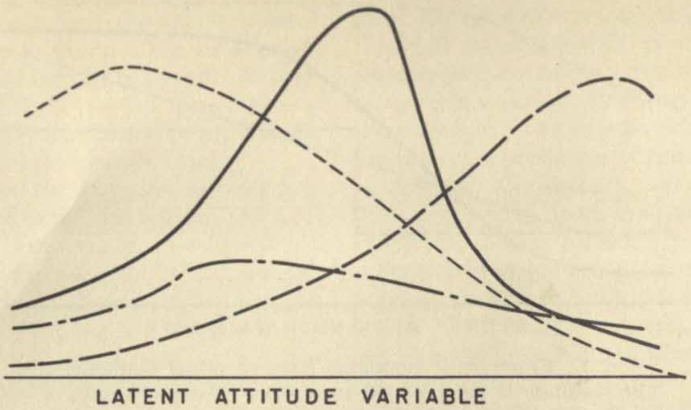


FIG. 2. Examples of nonmonotone item operating characteristics. The ordinate is the expected item score (the probability of a positive response, for dichotomous items).

question is an example. In our first nonmonotone item, if several short persons were thinking of a convertible Chevrolet, we might have found two maxima. Such items are of little value except perhaps in very special circumstances, and should normally be discarded or revised.

The distinction between monotone and nonmonotone items has frequently not been considered in scaling. However, it is particularly important, since Thurstone's scaling methods require nonmonotone items with operating characteristics having single maxima, while Guttman and Likert base their respective methods on monotone items.

QUESTIONNAIRES

A questionnaire is simply a collection of items. Usually the items are all of the same type. If several types of items are used, each type should be considered separately in the analysis unless the items can be shown to be sufficiently similar. The items on a questionnaire may or may not be of the same form, depending on which scaling method is to be used.

SCALING METHODS

Six major scaling techniques will be treated here: the judgment methods, the method of summated ratings, scalogram analysis, the unfolding technique, latent structure analysis, and rating methods. In the judgment methods the scale is developed from judges' ratings of the items. The method of summated ratings, scalogram analysis, the unfolding technique, and latent structure analysis develop scales from the data given by respondents, and hence may be called response methods. The rating methods use judges to rate the individuals.

JUDGMENT METHODS

The judgment methods of attitude measurement are all variants of a technique developed by Thurstone (1928a, b) and reported in detail by Thurstone and Chave (1929). The technique requires dichotomous, nonmonotone items having operating characteristics with single maxima. Although any type of nonmonotone items can be scaled by the judgment methods, opinion statements are usually used. Each statement represents a certain degree of favorableness toward a set of social objects, events, or ideas. Below are some representative items from Droba's (1930) scale for measuring attitude toward war:

> All nations should disarm immediately.
>
> Wars are justifiable only when waged in defense of weaker nations.
>
> The benefits of war outweigh its attendant evils.

There are two major aspects of the judgment methods. First, the items must be scaled; that is, each item must be given a scale value that represents its degree of favorableness toward the issue. Secondly, the respondents must be scored on the basis of their responses to the items.

Scaling the items. There are several methods of determining the scale values of the items. If

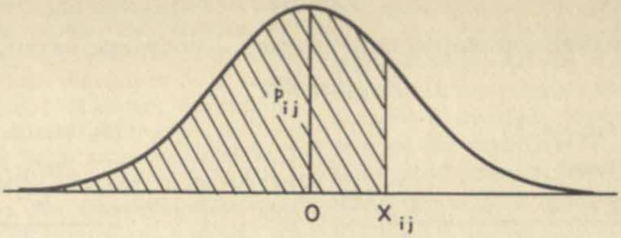


FIG. 3. The unit normal deviate transformation. The abscissa is measured in units of one standard deviation.

only a few items are to be scaled, the method of paired comparisons is appropriate (Thurstone, 1927; Guilford, 1936, Chapter VII). Each of several judges (preferably 25 or more) is presented with every possible pair of items, and is asked which of the pair is more favorable to the issue in question. For each pair of items we obtain the proportion of times one statement was judged to be more favorable than the other. In the mathematical model it is assumed that the perceived differences between stimuli have normal distributions. Specifically, the model states that

$$S_i - S_j = X_{ij}\,\sigma_{(i-j)}, \tag{1}$$

where S_i and S_j are the scale values of items i and j, $\sigma_{(i-j)}$ is the standard deviation of the hypothetical distribution of differences, and X_{ij} is the unit normal deviate corresponding to p_{ij}, the proportion of times item i is judged to be more favorable than item j, i.e.,

$$p_{ij} = \int_{-\infty}^{X_{ij}} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt. \tag{2}$$

The relation between X_{ij} and p_{ij} is illustrated in Fig. 3.

In the simplest case of the model, the so-called Case V, all $\sigma_{(i-j)}$ are assumed to be equal and are set equal to unity. In this case the model becomes

$$S_i - S_j = X_{ij}. \tag{3}$$

This simple model will usually suffice.

The data may or may not be consistent with the model. For a set of scale values, we can use equation (3) to compute an expected set of X_{ij}. These values can then be compared with the observed data. Since there are many more experimental data than there are scale values, some discrepancies between the expected X_{ij} and the observed X_{ij} are inevitable. Our problem is to find a set of scale values for which the discrepancies are minimum. Mosteller (1951a) has shown that Thurstone's method of estimation, described below, is a least-squares solution — the sum of squared discrepancies is minimum.

Mosteller (1951b) has also provided a chi-square test of the fit of the model. Both the expected X_{ij} and the observed X_{ij} are converted into proportions by means of the unit normal deviate transformation — equation (2). These proportions, labeled p''_{ij} for the expected proportions and p'_{ij} for the actual data, are transformed by the arc sin transformation

$$\theta_{ij} = \arc\sin\sqrt{p_{ij}}. \tag{4}$$

A table of this transformation is available in Snedecor (1946, p. 449). This gives us a set of θ'_{ij} and a set of θ''_{ij} which we use with equation (5) to compute chi square:

$$\chi^2 = \frac{\sum\limits_{i<j} (\theta'_{ij} - \theta''_{ij})^2}{821/n}. \tag{5}$$

Here n is the number of judgments on which each observed proportion is based, and the summation is over all values of i and j in which i is less than j, a total of $\frac{1}{2}k\,(k-1)$ terms, where k is the number of items. This chi-square value has $\frac{1}{2}(k-1)\,(k-2)$ degrees of freedom.

The first step in obtaining the scale values is to convert the observed proportions p_{ij} into X_{ij} by equation (2), using a table of the normal distribution. These X_{ij} are arranged in a two-way table, or data matrix, with a row for each item i and a column for each item j, as in Table 1.

Note that X_{ii} is logically zero, and that $X_{ij} = -X_{ji}$. If any X_{ij} is greater than $+2.0$, or less than -2.0, it should be rejected as unstable. If no values of X_{ij} have been rejected by this rule, the scale value of item i is the average of all entries in the ith column of the data matrix, as shown in Table 1 (b).

TABLE 1.

PAIRED COMPARISONS COMPUTATIONS — COMPLETE MATRIX

(4 ITEMS)

(a) Matrix of p_{ij}

		Item j 1	2	3	4
	1	—	.58	.83	.95
	2	.42	—	.80	.92
Item i	3	.17	.20	—	.65
	4	.05	.08	.35	—

(b) Matrix of X_{ij}

		Item j 1	2	3	4
	1	0	.20	.95	1.65
	2	− .20	0	.84	1.41
Item i	3	− .95	− .84	0	.39
	4	−1.65	−1.41	−.39	0
Column sums		−2.80	−2.05	1.40	3.45
Scale values		− .70	− .51	.35	.86

TABLE 2.

PAIRED COMPARISONS COMPUTATIONS — INCOMPLETE MATRIX

(6 ITEMS)

(a) Matrix of p_{ij}

		Item j 1	2	3	4	5	6
	1	—	.55	.75	.87	.98	1.00
	2	.45	—	.65	.79	.95	1.00
	3	.25	.35	—	.75	.91	.99
Item i	4	.13	.21	.25	—	.84	.92
	5	.02	.05	.09	.16	—	.78
	6	.00	.00	.01	.08	.22	—

(b) Matrix of X_{ij}

		Item j 1	2	3	4	5	6
	1	0	.13	.67	1.13	—	—
	2	− .13	0	.39	.81	1.65	—
	3	− .67	− .39	0	.67	1.34	—
Item i	4	−1.13	− .81	− .67	0	.99	1.41
	5	—	−1.65	−1.34	− .99	0	.77
	6	—	—	—	−1.41	− .77	0

(c) Matrix of successive differences

		Items 2-1	3-2	4-3	5-4	6-5
	1	.13	.54	.46	—	—
	2	.13	.39	.42	.84	—
	3	.28	.39	.67	.67	—
Item i	4	.32	.14	.67	.99	.42
	5	—	.31	.35	.99	.77
	6	—	—	—	.64	.77
Column sum		.86	1.77	2.57	4.13	1.96
Column average		.22	.35	.51	.83	.65

$S_1 = 0$, $S_2 = .22$, $S_3 = .57$, $S_4 = 1.08$, $S_5 = 1.91$, $S_6 = 2.56$

(d) Chi-square computation

ij	p'_{ij}	p''_{ij}	θ'_{ij}	θ''_{ij}	$\theta'_{ij} - \theta''_{ij}$
12	.55	.587	47.87	50.01	−2.14
13	.75	.716	60.00	57.80	2.20
14	.87	.860	68.87	68.03	.84
23	.65	.637	53.73	52.95	.78
24	.79	.805	62.72	63.79	−1.07
25	.95	.955	77.08	77.75	− .67
34	.75	.695	60.00	56.48	3.52
35	.91	.910	72.54	72.54	.00
45	.84	.797	66.42	63.22	3.20
46	.92	.931	73.57	74.77	−1.20
56	.78	.742	62.03	59.47	2.56

$$n = 100$$

$$\sum_{i<j} (\theta'_{ij} - \theta''_{ij})^2 = 42.95$$

$$\chi^2 = 5.23$$

$$\text{d.f.} = 6$$

$$.70 > p > .50$$

If some X_{ij} have been rejected, as indicated in Table 2 (a), then for each pair of successive columns of the data matrix, we subtract the entries in one column from the corresponding entries in the other column, and record these differences in a second matrix, as in Table 2 (b). Of course, no difference can be obtained whenever one of the values has been rejected. Since we are using successive columns, it may be possible to arrange the order of the columns so that there are very few impossible differences. The average of these differences for a pair of columns is the difference between the scale value for those two items. Thus, by arbitrarily setting $S_1 = 0$, the remaining scale values can be determined.

Note that both the zero point and the scale factor are arbitrary in the scale. We set the scale factor by letting all standard deviations equal unity. In the case of the complete matrix, the zero point was set by letting the average scale value be zero. In the case of the incomplete matrix, the zero point was set by letting $S_1 = 0$. Any other arbitrary values for the zero point and scale factor may be used.

The application of the chi-square test is illustrated in Table 2 (c).

The method of paired comparisons becomes unwieldy for large numbers of items, since each possible pair of items must be judged. For k items there are $\frac{1}{2}(k)$ $(k - 1)$ pairs; for 25 items there are 300 pairs. Several methods may be used as short cuts in this situation, including the methods of rank order, of equal-appearing intervals, and of successive intervals.

The method of rank order, or order of merit method (Guilford, 1936, Chapter VIII), is seldom used. In this method each judge places the items in rank order of favorableness toward the issue in question. The scale value of an item is the weighted or unweighted average of the ranks given to the item by a group of judges. Alternatively, a paired comparisons analysis can be used, following the procedure suggested by Thurstone (1931). For large numbers of items, the method of rank order becomes unwieldy, since the judge must make many decisions about the rank order of items which are nearly equal in degree of favorableness.

The method of equal-appearing intervals (Thurstone and Chave, 1929; Guilford, 1936, Chapter V) is the most widely used of the judgment methods. It is much simpler in application than the method of paired comparisons, since it requires only one judgment per item. Each judge is required to sort the items into a fixed number of categories (usually 7, 9, or 11) that are arranged in order of favorableness. The judge is usually asked to put items into categories in such a way that the intervals between categories are subjectively equal. That is, it should appear to him that the difference in degree of favorableness between the statements of any two adjacent categories is equivalent to the difference in favorableness between any other two adjacent categories. However, in this method there is no way of determining whether, in fact, the intervals are subjectively equal. The equality of intervals must remain as an assumption of the experimenter. It is not a particularly good assumption, since judges tend to make coarser discriminations in the mid-range of the scale than at the extremes. Hevner (1930) demonstrated this end effect empirically by constructing a scale by the method of equal-appearing intervals and also by paired comparisons. The relation between the two sets of scale values was approximately linear except at the extremes, where the equal-appearing interval scale was relatively compressed.

The method of successive intervals, first reported by Saffir (1937), was designed by Thurstone to overcome some of the difficulties inherent in the method of equal-appearing intervals. The method of successive intervals is recommended when there are too many items to be scaled feasibly by the method of paired comparisons (say 20 or more).

Each judge is required to sort the items into a fixed number of categories (usually 7, 9, or 11) spaced along a continuum of favorableness. The categories have a rank order on the continuum, but they are not necessarily equally spaced. The items in a given category are simply required to be less favorable than those in the adjacent category to the left and more favorable than those in the adjacent category to the right. When a number of judges have sorted the items, we can tabulate a judgment distribution for each item. This distribution shows the proportion of judges who placed the item in each category.

The mathematical model treats the categories as contiguous segments of the attitude continuum, separated by category boundaries. In the model it is assumed that the category boundaries can be located on the continuum in such a way that all the item judgment distributions are normal. The formal model is stated in terms of the cumulative judgment distributions. Let the M categories be numbered in order from least to most favorable. Let p_{jg} be the proportion of judges who placed item j in category g or any category of lower rank order. Let X_{jg} be the unit normal deviate [see equation (2) and Fig. 3] corresponding to p_{jg}. In the model, t_g is the boundary between category g and category $g + 1$. S_j is the mean of the judgment distribution for item j, and hence the scale value for item j, and σ_j is the standard deviation of item j. The model is illustrated in

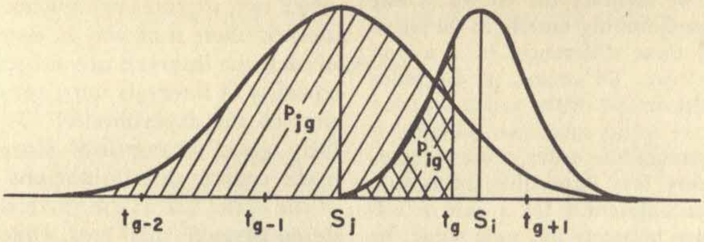


FIG. 4. The successive intervals model. Two item judgment distributions with different standard deviations are shown. The abscissa is the latent attitude variable, and is measured in arbitrary units.

Fig. 4. It states that

$$X_{jg} = \frac{t_g - S_j}{\sigma_j}. \qquad (6)$$

A graphical procedure for obtaining the values of t_g, S_j, and σ_j is presented by Saffir (1937). Mosier (1940a) described a short cut that is useful when a large number of items is involved. A least-squares solution has been developed by Gulliksen (1952), but its application is quite complicated.

Fortunately, it is usually adequate to assume that all standard deviations, σ_j, are equal, in which case they are set equal to unity. This model states that

$$X_{jg} = t_g - S_j. \qquad (7)$$

Procedures for obtaining the scale values and the category boundaries have been described by Saffir (1937), Attneave (1949), who uses the term "method of graded dichotomies," Garner and Hake (1951), who use the term "equal discriminability scale," and Edwards (1952). A least-squares method similar to these methods is described below.

The method of successive intervals is based on the assumption that the item judgment distributions are normal on the derived scale. It should be noted that this is not an assumption about the distribution of attitude scores of those who endorse the item, but rather an assumption about the distribution of judgments concerning the proper location of the item on the scale. This assumption must be checked in order to ascertain whether the model fits the data. The check may be accomplished graphically by plotting the observed p_{jg} versus the estimated t_g on normal probability paper. Straight lines should result, and if equal standard deviations have been assumed, these lines should have the same slope. Wide departures from these conditions are viewed with alarm. Edwards and Thurstone (1952) suggest obtaining the differences between the observed proportions and the corresponding proportions computed from the estimated parameters of the model. Although no statistical criterion is presented, the authors give some representative average discrepancies.

The first step in obtaining the scale values and category boundaries is to convert the observed proportions p_{jg} into X_{jg} by the unit normal deviate transformation, using a table of the normal distribution. The X_{jg} are arranged in a data matrix with a row for each item, and a column for each category. This is illustrated by an artificial example with 6 items and 5 categories in Table 3. All X_{jg} that are greater than $+2.0$ or less than -2.0 are rejected. If no values are rejected, the average value of each column is the category boundary represented by that column, while the average value of each row, when subtracted from the over-all average of the table, gives the scale value of the corresponding item.

If some values are rejected, then for adjacent columns we subtract the entries in one column from the corresponding entries in the other column, and the differences are recorded in a column of a second matrix. Of course, wherever a value has been rejected, no difference is possible. The average of these differences is the difference between the category boundaries represented by the two columns. By assigning $t_1 = 0$, the remaining category boundaries can be computed, as shown in Table 4. Then a fourth matrix is constructed by subtracting the computed value of t_g from each entry in the gth row of the X_{jg} matrix, as shown in Table 4 (d). The average value of a row of this matrix is the scale value of the corresponding item.

Since the scale factor and zero point are arbitrary, the scale factor has been set by letting all standard deviations equal unity. In the case of the complete matrix, the zero point was fixed by letting the average scale value be zero. In the case of the incomplete matrix, the zero point was fixed by setting $t_1 = 0$. Any other arbitrary scale factor and zero point may be used.

TABLE 3.

SUCCESSIVE INTERVALS COMPUTATIONS — COMPLETE MATRIX

(5 CATEGORIES, 6 ITEMS)

(a) Matrix of p_{ij}

		Category				
		1	2	3	4	5
	1	.40	.55	.86	.96	1.00
	2	.28	.36	.71	.90	1.00
	3	.15	.30	.65	.81	1.00
Item j	4	.16	.25	.54	.82	1.00
	5	.09	.21	.50	.74	1.00
	6	.03	.08	.37	.57	1.00

(b) Matrix of X_{jg}

		Category boundary g				Row	Row	Scale
		1	2	3	4	sum	average	value
	1	− .25	.13	1.08	1.75	2.71	.68	− .78
	2	− .58	− .36	.55	1.28	.89	.22	− .32
	3	−1.04	− .52	.39	.88	− .29	− .07	− .03
Item j	4	− .99	− .67	.10	.92	− .64	− .16	.06
	5	−1.34	− .81	.00	.64	−1.51	− .38	.28
	6	−1.88	−1.41	− .33	.18	−3.44	− .86	.76
Column sum		−6.08	−3.64	1.79	5.65	−2.28		
Category boundary		−1.01	− .61	.30	.94		− .10	

TABLE 4.

SUCCESSIVE INTERVALS COMPUTATIONS — INCOMPLETE MATRIX

(5 CATEGORIES, 6 ITEMS)

(a) Matrix of p_{jg}

		Category				
		1	2	3	4	5
	1	.44	.70	.94	.99	1.00
	2	.16	.45	.81	.93	1.00
	3	.04	.24	.52	.78	1.00
Item j	4	.04	.16	.40	.69	1.00
	5	.01	.08	.32	.57	1.00
	6	.00	.02	.14	.36	1.00

(b) Matrix of X_{jg}

		Category boundary g			
		1	2	3	4
	1	− .15	.52	1.56	—
	2	− .99	− .13	.88	1.48
	3	−1.75	− .71	.05	.77
Item j	4	−1.75	− .99	− .25	.50
	5	—	−1.41	− .47	.18
	6	—	—	−1.08	− .36

(c) Matrix of differences

		2-1	3-2	4-3
	1	.67	1.04	—
	2	.86	1.01	.60
	3	1.04	.76	.72
Item j	4	.76	.74	.75
	5	—	.94	.65
	6	—	—	.72
Column sum		3.33	4.49	3.44
Column average		.83	.90	.69

$t_1 = 0, t_2 = .83, t_3 = 1.73, \; t_4 = 2.42$

(d) Matrix of $t_g - X_{jg}$

		Category boundary				Row	Scale
		1	2	3	4	sum	value
	1	.15	.31	.17	—	.63	.21
	2	.99	.96	.85	.94	3.74	.94
	3	1.75	1.54	1.68	1.65	6.62	1.66
Item j	4	1.75	1.82	1.98	1.92	7.47	1.87
	5	—	2.24	2.20	2.24	6.68	2.23
	6	—	—	2.81	2.78	5.59	2.80

The methods of paired comparisons and successive intervals yield interval scales in which the metric is obtained from the variability of judgments about the items. If items were always judged consistently in paired comparisons, so that the proportions were all either zero or unity, the mathematical model of paired comparisons could not be used. Identical sortings from all judges would render the method of successive intervals equally useless. Happily, judges seldom agree. There is usually sufficient "error" variance to allow a scale to be constructed. The metric in these scales is essentially a function of the degree of confusion, or overlap, between items. The difference of the scale values of two items may be used to make a precise probability statement about the degree of overlap. The scales are based on variability, but it is a remarkably predictable variability.

Scoring the respondents. When a scale of items has been established by one of the judgment methods, a rule for scoring the respondents must be adopted. The usual scoring procedure is to give each person a score which is the median of the scale values of the items which he endorses. The mean scale value could be used, but it is felt that the median is less sensitive to the particular set of scale values in the scale. Lorge (1939) found very little difference between mean scores and median scores. By either scoring method, the respondents are placed on the scale of favorableness toward the issue being studied.

Item selection. Usually, many more items are scaled than will be used on the final questionnaire. Ambiguous or irrelevant items can be discarded when the items are picked for the final form of the scale. If the judgment distribution for an item has a large standard deviation, the item is presumed to be ambiguous, since it apparently means different things to different judges. In the extreme case of ambiguity, the judgment distribution might be spread uniformly over the categories. Items that are ambiguous according to this subjective criterion should be eliminated before the successive intervals scale is computed, so that the assumption of equal standard deviations becomes tenable. The simple estimation procedures described above can then be used.

Irrelevant items cannot easily be detected on the basis of judgments. If an item is almost always judged to be in the middle category, in the methods of successive, or equal-appearing, intervals, then it may be irrelevant, but it may also be a good neutral item. Thurstone (Thurstone and Chave, 1929) proposed a criterion of irrelevance which can be used when the questionnaire has been administered to a group of respondents. For each item in turn, the individuals who responded positively to a given item are selected; the popularity of each of the other items is determined for this group of individuals. These popularities are compared with the scale values of the items. Items with scale values near the given item should be very popular, while items far removed on the scale should have low popularity for this group. If the popularities are independent of the scale values, then some factor other than the latent attitude variable is determining the responses to this item.

Specifically, the relevance of item i is determined by plotting the ratio p_{ij}/p_j versus S_j. The divisor p_j is used to correct for the different popularities of items in the total group.

A similar technique can be used to get a rough picture of the operating characteristics of the items. In this case, individuals of nearly equal attitude scores are grouped together. The proportion of individuals who endorse the item can be determined for each group, and can be plotted as a function of the mean attitude score of the group. An acceptable operating characteristic will have a maximum near the scale value of the item, and will decrease toward zero on either side of this maximum.

Items having unacceptable operating characteristics, or giving indications of irrelevance, should be rejected. Although these criteria are subjective, they allow the investigator to compare items, and then to choose the most relevant items and the items with the best operating characteristics.

When the poor items have been discarded, the investigator may still have more items than he needs for the final scale. In this case, he should select a set of items with scale values spread evenly along the continuum. If particularly good discrimination is desired in some portion of the continuum, then relatively more items may be selected with scale values in that range. If enough items are available, parallel forms of the scale can be produced by using items with matched scale values.

Criticisms of the judgment methods. Several authors have criticized the central, or neutral, items on Thurstone-type attitude scales for failing to discriminate high and low groups on the attitude continuum. This criticism stems from a lack of understanding of nonmonotone items. An ideal item characteristic curve for a neutral item would have zero probability of response at both extremes. In fact, if a neutral item does discriminate between high and low groups, it is not a very good neutral item.

According to Edwards (1946), items with scale values in the middle range often have higher judgment variances than other items. This may be due in part to the curtailed distributions of extreme items in the method of equal-appearing intervals. But it may also be due in part to ambiguity. Edwards also states that neutral items are more likely to be irrelevant according to the criterion of relevance. These troubles emphasize the difficulty of writing neutral items. No matter how hard one tries to avoid it, such items have a tendency to be "double-barreled."

The second major criticism is that the opinions of the judges might influence their ratings. Advocates of the judgment methods have denied this, stating that the only requirement for a judge is that he be able to read. Studies by Hinckley (1932), Ferguson (1935), and Pintner and Forlano (1937) all strongly support the contention that the attitudes of the judges do not affect the resulting scale. In these studies, the judges were separated into high and low groups on the basis of attitude scores. Scales were derived from each group of judges; the scale values correlated better than .95. Recently, Hovland and Sherif (1952) have criticized these studies for not using judges with extreme attitudes. Using Hinckley's scale of attitude toward Negroes, they found that judges with extreme attitutdes do yield quite different scales. Hinckley had rejected all judges who placed more than 25% of the statements in a single category, since he felt that they were being "careless." Hovland and Sherif find that these "careless" judges have very extreme attitudes and deep involvement, and that the single category that they favor is usually an extreme category. Scales obtained from Negro judges and from pro-Negro white subjects were quite different from scales obtained from "average" white subjects. The authors point out that this result is consistent with other facts of social psychology. Many studies indicate that attitudes, desires, and needs may influence perception, judgment, and other behavior. Kelley, Hovland, Schwartz, and Abelson (1953) have investigated the problem of biased judges further. They conclude that only in cases of extreme bias will there be a serious effect on the scale. They point out that Hovland and Sherif used the method of equal-appearing intervals, and that by using the method of successive intervals, the differences between the scales can be reduced substantially. The method of paired comparisons nearly eliminates the differences.

It appears that we can expect the bias effect to be small or negligible in most cases. However, it seems clear that the problem cannot be overlooked. More research is needed to determine the extent of the bias, and the conditions under which it occurs.

THE METHOD OF SUMMATED RATINGS

The Likert-type scale. A few years after the Thurstone scaling methods were introduced, Likert (1932) published a technique for the measurement of attitudes which was similar to methods in use in the mental testing field. In Likert's method of summated ratings, five categories of response are provided for each item: *strongly approve, approve, undecided, disapprove,* and *strongly disapprove.* These categories are scored 5, 4, 3, 2, and 1 respectively. If an item is cast in the negative sense, with *strongly approve* indicating an unfavorable attitude, then the category scoring is reversed. An individual's scale score is the sum of his scores on the items. (Likert originally scored the response categories of an item on the basis of a normal curve transformation of the proportions of respondents choosing each category. This "sigma scoring" method was immediately abandoned when it was discovered that the simpler 5 4 3 2 1 scheme yielded scores which correlated better than .99 with the sigma scores.)

The items on a Likert-type scale should have operating characteristics that are monotonically increasing functions of the latent attitude continuum. That is, the more favorable a person's attitude, the higher his expected score for the item would be. The items are evaluated by administering them to a group of respondents. The relations of the item score to the total score on the set of items is an index of the discriminating power of the item. Items with high discriminating power are selected for the final form of the questionnaire.

One widely used index of item discrimination is based on the differences in item scores for subgroups with high and low total scores, respectively. From the distribution of total scores, the highest 25% (or 27% – the exact percentage is of minor importance) of the population is selected for the high group; the lowest 25% forms the low group. For each group the mean, M, and the variance, σ^2, of the item scores are computed for each item separately. The index used is usually the critical ratio:

$$\text{C.R.} = \frac{M_{high} - M_{low}}{\sqrt{\dfrac{\sigma^2_{high} + \sigma^2_{low}}{n}}},$$

where n is the number of respondents in each group. This index evaluates the mean difference relative to the item score variances. It should be noted that the proper index is the critical ratio, not the probability of statistical significance.

Another useful index is the correlation between item scores and total scores. In many ways, this index is superior to the critical ratio, since it makes direct use of the scores of all respondents, rather than just the scores of the two subgroups.

These procedures of item analysis are very similar to mental test methods, although mental test items are usually scored 0 or 1. Gulliksen (1950b) has provided thorough discussions of problems such as test scoring, category weighting, and item analysis that are common to attitude measurement and mental testing. Davis (1951, 1952) has reviewed a large number of item analysis techniques, many of which are applicable to Likert-type scales.

The single common factor model. Likert did not provide a mathematical model for his scaling method. However, the scoring technique of the Likert scaling technique implies a model that we shall call the single common factor model. In this model, which is derived directly from factor analysis, the intercorrelations of the items are assumed to be due to a single common factor to which the items are all mutually related. Since the latent attitude dimension is defined by the mutual interrelations or consistency among items, the items are assumed to be interrelated only by virtue of the latent factor. With the effect of this factor held constant, the items are unrelated, or independent. This is the basic idea of both factor analysis and latent structure analysis. Mathematically, the item score is considered to be the weighted sum of two independent factors: the general attitude factor, and a factor specific to the item.

The model carries with it a justification of the scoring procedure, and in this sense is implied by the scoring procedure. As the number of items increases, the linear correlation of the total score with the single common factor approaches unity. For 25 items, a typical value of this correlation would be about .90. The reason for this near-perfect linear correlation is that as we add more and more items, the effects of the general factor cumulate very fast, since it is represented in each item, while the effects of the unique factors cumulate more slowly, since these factors are assumed to be mutually independent, and hence tend to cancel each other out, to some extent. The mathematical analysis will not be given here, but it parallels the discussions of weighting given by Wilks (1938) and by Gulliksen (1950b, Chapter 20).

The model stated here assumes a linear regression of item score on the common factor. This assumption leads to a linear regression of total score on the common factor, and to a perfect linear correlation between total score and factor as the number of items becomes infinite. Lord (1953) has pointed out that this model cannot be precisely correct, since it would theoretically allow any item score to occur, whereas actually the item score must be between 1 and 5. However, the restriction is probably not important.

A different model based on the single common factor idea has been developed recently by Lord (1952). He deals with dichotomous items, and assumes that the item operating characteristics are normal ogives, following Tucker (1946) and others. He shows that the distribution of scale scores depends on the item characteristics and may be quite different from the latent variable distribution. According to the model, there is a perfect curvilinear correlation between scale score and latent factor, as the number of items becomes infinite. The regression of the latent factor on the test score has an ogive shape.

One way to verify the single common factor model would be to perform a factor analysis of the item intercorrelations. Product-moment correlations would be used for the linear model, while tetrachoric correlations are appropriate for the ogive model. Certain precautions must be observed when factoring items, since the analysis is very sensitive to the item response distributions. Ferguson (1941) and Wherry and Gaylord (1944) have discussed this problem for dichotomous items. The difficulty is less crucial for five-category items, but it must be taken into account.

In practice, high item-scale correlations are sometimes taken as presumptive evidence of a single common factor. This is not necessarily valid evidence, since two or three equally potent factors might be represented by the items. An alternative test of the single factor model is to determine the homogeneity of the scale, since a homogeneous scale is a single factor scale.

The two factor models presented show that the relation of the score distribution to the single common factor depends on the particular model used to define the factor. It can also be shown that the score distribution depends on the operating characteristics of the particular items in the questionnaire. The investigator can freely manipulate the distribution by selecting the items carefully. Since the scoring method is insensitive to the characteristics of the particular items, and since it does not depend on the factor model, it has no unique rational metric, and must be regarded as yielding an ordinal scale.

There is a sense in which a Likert-type scale has a metric. For any given questionnaire, the score distribution will be fixed for the general population. Then a metric is possible in the sense that a difference between scores is related to the proportion of people having intermediate scores. Furthermore, percentile scores can be used. These scores are free from the troubles of the raw total scores. The percentile scores should be constant, except for unreliability, no matter how the raw score distribution is manipulated by item selection.

Maximum variance scores. The possibility of weighting the items differentially instead of simply adding the unweighted item scores has been investigated by several authors. It is well recognized that any weighted combination of items is very highly correlated with any other weighted combination, if corresponding weights have the same sign, and if the number of items is large. Weighting becomes a problem when the number of items is small. Gulliksen (1950b) has given a comprehensive discussion of item weights. We shall mention only one approach to the problem.

Horst (1935, 1936), Wilks (1938), and Edgerton and Kolbe (1936) suggested a method for combining scores so that the resulting score distribution would have maximum variance, under the restriction that the sum of squared weights is unity. The same concept was used by Guttman (1941) to develop weights for each item response separately. The result is known mathematically as the first principal component of the matrix of covariances of the variables to be weighted. The use of this scoring method does not change the arbitrary nature of the score distribution, since it will be affected by item selection in the same way as the unweighted score distribution.

A practical method for obtaining approximations to maximum variance scores is discussed by Niven (1953), who calls the technique *the method of reciprocal averages.*

SCALOGRAM ANALYSIS

A new departure in attitude measurement was provided by Guttman (1944), who proposed a nonmetric method for scaling monotone attitude items. In a Guttman scale the items have a special cumulative property. For example, a person who responds positively to the third item on the scale is almost sure to have responded positively to the first and second items.

Similar notions had been used in earlier studies, but not as a general method. One of the earliest examples of measurement in social psychology is the social distance scale of Bogardus (1925). He used a set of seven items denoting varying degrees of social acceptability. The acceptability, or social distance, of a nationality group was ascertained by noting which of the items were endorsed by a respondent. The items included admission to this country as a visitor only, admission to citizenship, to equal employment, admission as neighbors, as personal chums, and as relatives by marriage. Anyone willing to admit members of the group as neighbors would almost certainly be willing to indicate acceptance of items earlier on the list.

A second instance of the use of cumulative items was reported by May and Hartshorne (1926). They used six tests of cheating that could be arranged in order from the test on which fewest children cheated to the test on which most children cheated, or the most tempting test. A child who would cheat on one of the tests would also cheat on the more tempting tests.

However, the implication of these early studies for a general method of scaling were largely ignored. It was Guttman who advocated the use of such cumulative items as a basis for a formal method of scaling. His development of scalogram analysis stimulated widespread interest in the method. A recent and comprehensive account of the method will be found in Guttman (1950b).

The model. Although scalogram analysis is applicable to items with any number of response alternatives, the scale can most easily be described for dichotomous items. The basic idea is that the items can be arranged in an order so that an individual who agrees with, or responds positively to, any particular item also responds positively to all items of lower rank order. The rank order of the items is the scale of items; the scale of persons is very similar, people being arranged in order according to the highest rank order of items checked, which is equivalent to the number of positive responses in a perfect scale.

As an example of a perfect Guttman scale, consider a questionnaire concerning height. The items might be: (1) I am more than 5′0″ tall, (2) I am more than 5′3″ tall, (3) I am more than 5′6″ tall, etc. Very little inconsistency of response would be expected. The operating characteristic of a perfect Guttman item would indicate a response probability of zero up to a certain point on the continuum,

TABLE 5.

SCALOGRAM DIAGRAM FOR THREE DICHOTOMOUS ITEMS.

+ indicates a positive response, — indicates a negative response.

		Items		
		3	2	1
Individuals (scale types)	I	+	+	+
	II	—	+	+
	III	—	—	+
	IV	—	—	—

and from that point on a response probability of unity. In Table 5 is shown a diagram of a perfect Guttman scale of three items. In this scale an individual who gives the positive response to item 3 also gives the positive response to items 1 and 2. On the other hand, a person who gives the positive response to item 2 may or may not also agree with item 3; it is certain that he agrees with item 1.

When the items have been arranged in order, it is also clear that the persons have been divided into a number of groups, or *scale types,* according to their pattern of item responses. In the example cited there are four typical response patterns: (+ + +), (— + +), (— — +), (— — —). The nonscale response patterns which do not occur in a perfect scale are (+ — +), (+ + —), (— + —), and (+ — —). The score of an individual is determined by assigning the integers in order to the scale types, viz., 3, 2, 1, 0. In the case of dichotomous items, the score is equivalent to the number of positive responses.

For a perfect Guttman scale, the order of the items is the same as the order of their popularities. Also, every 2 × 2 contingency table showing the relation between two items has an empty cell, since no one who responds negatively to the more popular item will respond positively to the less popular item. This is an important mathematical property of the perfect Guttman scale for dichotomous items: if $p_i \geqq p_j$, then $p_{ij} = p_j$, where p_{ij} is the proportion who respond positively to both items i and j.

In the case of items with several response categories, the scale pattern becomes slightly more complex. Here each response category of an item is treated as a separate item; it is necessary to find an order for these response categories. The precise definition of a scale is given by Guttman (1950b, p. 62): "We shall call a set of items of common content a scale if a person with a higher rank than another person is just as high or higher on every item than the other person." Items with several response categories that have a logical order can be reduced to dichotomous items for the analysis. A single item with five response alternatives would be replaced by four dichotomous items. The positive responses to these four items would be equated with responses in category 1, categories 1 or 2, categories 1, 2, or 3, and categories 1, 2, 3, or 4, respectively. By this simple expedient the dichotomous model can be made to include the nondichotomous case as well.

When a set of items satisfies the scalogram model, the scale is said to be unidimensional. The model can be extended to the multidimensional case. Bennett (1952) considered the conjunctive and disjunctive models for many dimensions. Each individual has a position on each of several attitude dimensions. Each item has a breaking point on each dimension. In the conjunctive model, a person will respond favorably to an item only if his position is above the item's breaking point on all the dimensions. In the disjunctive model he will respond positively if he is above the item's breaking point on any one of the dimensions. Bennett (1951) has also developed methods for ascertaining the number of dimensions necessary to account for the observed item responses.

The scaling technique. In the practical business of attitude measurement, some of the nonscale response patterns are bound to occur. Ford (1950) has pointed out that the prototype of a cumulative set of attitude items is not "Are you taller than 5′0″?" etc., but rather, "Are you taller than a table?" "Are you taller than the head of a pony?" "Are you taller than a good-sized bookcase?" We do not expect these items to form a perfect scalogram. No matter what set of ordered response patterns is selected as the scale types, there will be some errors. The purpose of the scaling technique is to select a set of scale types so that the errors will be minimized.

In a perfect Guttman scale, an individual's score bears a one-to-one relation with his response pattern. That is, his item responses are completely specified by his score. From the score his item responses can be perfectly reproduced. For a nonperfect scale an individual receives the score of the scale type which is most nearly like his response pattern. That is, a score is assigned so that his response pattern is as nearly reproducible as possible.

Consider a scale of five dichotomous items, with scale types (+ + + + +), (+ + + + —), (+ + + — —), (+ + — — —), (+ — — — —), and (— — — — —). A person with response pattern (+ + — + +) would be classed in the (+ + + + +) category, and one error counted. The response pattern (+ + — — +) would be classified with (+ + — — —) and one error counted. The response pattern (+ — + — +) could be classed with (+ + + + +), or (+ + + — —), or (+ — — — —) for two errors. Henry (1952) proposed that in ambiguous cases, where several classifications are possible with the same error, the dilemma could be resolved by choosing the most popular of the alternative patterns. Borgatta and Hays (1952) suggest using the "average" of the alternatives. In the above examples the "average" of no minuses, three minuses, and four minuses is presumably two minuses, so the proper scoring of (+ — + — +) would be (+ + + — —). The argument is based on the Lazarsfeld latent distance model (*cf. infra*). (In practice, the number of positive responses may be used as the score, since it correlates highly with the more complex scoring methods.)

The percentage of errors of reproducing scale patterns is the basic measure of error. The quantity $1 - \dfrac{\text{number of errors}}{\text{number of item responses}}$ is called the coefficient of reproducibility, which we shall symbolize by Rep. Rep is the percentage of item responses which are correctly reproduced. The practical scaling problem is to arrange the items so that the resulting scale types yield the maximum Rep.

Several methods have been described for selecting the scale types. Most of these methods are based on the extended scalogram diagram, which is essentially the matrix of item responses, with a row for each individual and a column for each item response category. An extended diagram for a perfect scale of three dichotomous items is shown in Table 6.

TABLE 6.

EXTENDED SCALOGRAM DIAGRAM FOR THREE DICHOTOMOUS ITEMS.

X indicates that the individual responded in the item response category.
+ indicates the positive response category.
— indicates the negative response category.

		Item response categories					
		3+	2+	1+	3–	2–	1–
	I	*X*	*X*	*X*			
Individuals (scale types)	II		*X*	*X*	*X*		
	III			*X*	*X*	*X*	
	IV				*X*	*X*	*X*

Note that the last three columns are completely redundant. Note also that the responses form a parallelogram. Procedures are described for rearranging the rows and also the columns, as well as for combining response categories and perhaps even reversing the scoring of some items, so that a response parallelogram is achieved with few deviations. These methods are not automatic but require keen judgment concerning the kind of sorting likely to pay off. Suchman (1950a) describes the scalogram board procedure, which represents the response matrix by buckshot placed in small indentations in a set of removable wooden slats. The sorting is accomplished by interchanging these slats in their frame. Since both horizontal and vertical sorting is desired, two matching scalogram boards are needed. Patience and a steady hand are also helpful. Methods for tabulating the scalogram diagram on IBM equipment have been described by Noland (1945), Ford (1950), and Kahn and Bodine (1951). Paper and pencil methods have been described by Guttman (1947) and Marder (1952).

None of these methods is elegant or compact. Each one attempts to evaluate the complete data matrix without the aid of summary statistics. For large numbers of respondents, the task is overwhelming. It has been suggested that the scale be established on the basis of a sample of 100 cases. Since the statistical theory of the random fluctuation in reproducibility is unknown, this seems to be a dangerous extravagance. Moreover, it is difficult to deal with more than 10-20 items with these procedures.

The procedure recommended by the present author is simple. First, dichotomize all items by combining response categories or by defining additional items (see above). Second, order these items according to their popularities, with the least popular having the lowest rank. Goodenough (1944) and others have suggested this ranking procedure. Third, estimate Rep as follows: for each item, tabulate a_i, the proportion of cases giving the positive response to that item and the negative response to the item of next higher rank; add these a_j (with k items there are $k - 1$ a_j's), divide by k, and subtract from unity. This always gives an overestimate of Rep, but the error is slight. A validation of this estimate as well as a method for improving it by additional tabulations has been given by Green (1953).

An alternative useful method, called *H*-technique, for improving cumulative scales has been suggested by Stouffer, Borgatta, Hays, and Henry (1952). The details of their method will not be given here, but the basic notions will be outlined. All items are dichotomized by combining response categories. Items with similar popularities are grouped together; about four such groups are expected. The items within each group are combined to form a new contrived dichotomous item. For example, if there are three items in a group, a positive response on the contrived item might be defined as the occurrence of a positive response on any two or all three true items, while the negative response would be equated with a positive response on none or only one of the true items. Finally, the contrived items would be used to form a scalogram. (The method of the preceding paragraph can be used from this point.) The authors show that reproducibility will be enhanced by *H*-technique.

Criteria for a Guttman scale. Guttman has provided criteria for determining whether a set of items is acceptable as a scale, in his sense of the term. The major necessary condition is that the reproducibility, Rep, be at least .90. However, this is not a sufficient condition; additional criteria must also be met.

The need for auxiliary criteria arises from the unfortunate mathematical properties of Rep. Festinger (1947), Guttman (1950b), and others have pointed out that the size of Rep depends on several factors. The value of Rep for the scale is the average of the values of Rep for the items in the scale. The reproducibility of an item cannot be less than the popularity of the modal category, since the scale can do no worse than assign the most popular response in all cases. Thus the average popularity of the most popular response to each item is a lower bound of Rep. Items with high modal popularity will cause Rep to be spuriously high.

A second difficulty is that the expected value of Rep, if the items are actually independent, is often high. This can be illustrated for dichotomous items. Consider four independent items with marginals, p_i, of .2, .4, .6, and .8. The average popularity of the modal categories is $1/4\ (.8 + .6 + .6 + .8) = .7$. We can compute the expected proportion of people in each response category; e.g., $(+ + - +)$ has $.2 \times .4 \times .4 \times .8 = .0256$ or 2.56% of the total population. This proportion is assigned one error. Proceeding in this way, we compute the expected Rep to be .90. A set of nine items with marginals .1, .2, .3, .4, .5, .6, .7, .8, .9 has an expected Rep of .83, assuming independence of items. Festinger (1947) suggests that the expected Rep, under the assumption of independence, should be reported whenever an actual Rep is reported. Many Reps appear to indicate much less response consistency when evaluated in this way.

The auxiliary criteria are as follows:

1. There should be a good spread of item marginals, although few, if any, items should have marginals less than .2 or greater than .8.
2. Each response category should have more nonerror than error. Otherwise the category does not correlate with the scale.
3. If the items are dichotomous, at least ten items should be used.
4. Errors should be random. If any particular error occurs much more frequently than others, it is evidence of a *nonscale type*.

If these criteria are all met, but Rep is less than .90, the items are called a *quasi-scale*. Since errors are random in a quasi-scale, it has one property in common with an acceptable Guttman scale, namely, the curvilinear correlation (eta) of the scale scores with any outside variable is equal to the multiple correlation of the items with that variable. That is, if the scale score is held constant, the items are not correlated with each other or with any other variable. Actually this property seems to be the criterion for random errors in a Guttman scale.

The basic difficulty in determining whether a set of items forms a scale is that the Guttman model is not a probability model, so the theory of the statistical fluctuations to be expected is obscure, and must be based on a complex set of interdependent criteria. The latent distance model, discussed below, provides a probability model for a Guttman scale, and hence overcomes many of these difficulties.

If a set of items does not scale, the possibility exists of rejecting one or two poor items, and then achieving a scale. Guttman is chary of this procedure, preferring to say that the universe is not scalable. However, it seems possible to have perfectly good items with the wrong form for a Guttman scale. To this author, the possibility of rejecting items seems to be a necessary part of any method of attitude measurement.

Kriedt and Clark (1949) used scalogram analysis to select a reproducible scale. They complain that the method is cumbersome, due to the multiple criteria. They found that the resulting scale was less reliable than a scale constructed by item analysis methods from the same initial set of items.

Edwards and Kilpatrick (1948a, b) have suggested a method of item selection for a Guttman scale that uses judges and also item analysis techniques. A large set of dichotomous monotone items is assembled. Judges are asked to sort these items into categories according to degree of favorableness. Note that the judge's task is somewhat different than in the judgment methods where nonmonotone items are used. Here, the judge is probably rating according to the expected proportion of positive responses to the item. Items with a high judgment variance are rejected as ambiguous. The remaining items are scored 1 for a positive response, 0 for a negative response, and a total score is obtained. An item analysis is performed, as in the method of summated ratings. Nondiscriminating items are rejected, and the remainder are subjected to a scalogram analysis.

Loevinger (1947, 1948) has developed a theory of homogeneous tests that is equivalent to scalogram analysis. Her definition of homogeneity is identical to Guttman's concept of a perfect scale. Her coefficient of homogeneity, corresponding to Rep, is the ratio of the average inter-item covariance to the average covariance of a perfectly homogeneous set of items with marginals equal to those of the observed items. This index has some of the difficulties inherent in Rep, but it may prove to be superior; not enough is known at present to judge.

Intensity and a zero point. Guttman, Suchman, and their followers have suggested a rational method for obtaining a zero point on the attitude scale. Although this development is in terms of scalogram analysis, the method is applicable to any attitude scale. The basis for the development is that the intensity with which an attitude is held seems to be related to the attitude score by a U-shaped function, with people at either extreme of the attitude scale having intense attitudes, and people in the center being relatively apathetic. The attitude zero point is defined as that attitude score for which the intensity is minimum. That is, persons with an attitude score of zero don't care very much, one way or the other. Thus, if an independent measure of intensity is available, it is possible to define the zero point of the attitude objectively. An alternative procedure that may be necessary in practice is to define a region of indifference. This seems preferable when the intensity function does not have a sharply defined minimum.

Katz (1944), Cantril (1946), and Suchman (1950b) have suggested methods for measuring intensity. An intensity question can be asked after each attitude question. The intensity question might be, "How strongly do you feel about this?" with alternatives *very strongly, fairly strongly,* and *don't care.* Another possibility is the question, "How sure are you that your opinion is right?" with alternatives *very sure, fairly sure,* and *not sure.* Still another possibility is, "How much does this question mean to you personally?" Self-rating scales have been used, such as a thermometer on which the respondent indicates the intensity of his attitude. It has been suggested that if Likert-type alternatives have been allowed for each item, then one can lump *strongly agree* and *agree* as the positive response, the other categories being the negative response. Also, *strongly agree* and *strongly disagree* can be lumped as strong intensity, *approve* and *disapprove* as medium intensity, and *undecided* as low intensity. Since the attitude and intensity variables thus defined are not independent, it would seem better to ask two separate questions. Postman and Zimmerman (1945) found that decision time was correlated with intensity. They asked the subject to respond to each item on a ten-point scale from *strongly agree* to *strongly disagree.* Decision times were highest for items rated in the center of the scale.

When intensity questions are available, intensity may be scaled. Suchman suggests that the scalogram method of scaling be used for intensity, although he finds that intensity is usually a quasi-scale.

The definition of a zero point on an ordinal or interval scale is an arbitrary matter. In general, one point is logically as good or as bad as another. It seems reasonable, however, to require that the zero point should not depend on the particular set of items used. The definition of a zero point as the attitude score of minimum intensity, if adopted, would be a convention. Any justification must be on empirical grounds.

If it is desired to dichotomize the population into *pro* and *con* on an issue, the split might well be made at the region of minimum intensity. It is possible that one group will behave quite differently from the other. Suchman (1950b) has shown that two scales with biased selections of items can be compared by equating regions of indifference. This is good evidence of the utility of the intensity convention for a zero point.

Metric and components. To provide a metric for his scale, Guttman (1941, 1950c) proposed a criterion of scoring so that persons giving a certain response to an item would have scores as similar as possible among themselves and as different as possible from persons giving other responses to the item. This criterion is mathematically equivalent to the criterion for maximum variance scores discussed above.

Maximum variance scores do not play a major role in scalogram analysis, since the simple ranks of response categories are adequate for most purposes. However, the mathematical analysis is interesting, since it leads to several sets of scores, all of which satisfy the equations. These sets of scores are orthogonal, i.e., uncorrelated. The first set of scores has maximum variance. The scores in the second set have maximum variance under the restriction that they are uncorrelated with the first set. Scores in the third set have maximum variance under the restriction that they are uncorrelated with the first and second sets, etc. Mathematically, these sets of scores are known as principal components. It turns out that the first principal component has the largest variance, the second principal component has the second largest variance, etc. If there are M possible scale scores, i.e., M perfect scale types, then there are M principal components. The existence of these components is a straightforward mathematical property of the items and the mathematical problem of maximizing the score variance.

Guttman has shown that, for a perfect scale, if the values of the various sets of scores are plotted as functions of the original scale scores, the first principal component is a monotonic function of the scale scores. For this reason it can be used as a metric for the system. Moreover, the second principal component is U-shaped, having one inflection point; the third principal component has two inflection points; the fourth has three inflection points, etc.

Since the second component and the intensity function are both U-shaped, Guttman has equated the two functions, saying that intensity *is* the second principal component. The identification is logical only if it is taken to mean that intensity is related to the scales in roughly the same way as the second component. The similarity of the two functions seems to be quite accidental; intensity is not in any way more fundamental or meaningful because of it. Guttman believes that it will be possible to find other attributes of attitude which behave in roughly the same way as the third, fourth, and higher components. He has already identified the third component as "closure" (Guttman, 1950d).

The present author agrees that it is profitable to investigate the correlates of attitude. There may be many interesting characteristics of attitude, such as clarity, specificity, intensity, closure, involution, etc. Perhaps some of these characteristics will have several inflection points, and hence correspond with one of the higher components. If so, the correspondence will be entirely fortuitous. Whether or not such functions can be demonstrated reliably, and whether they will increase our understanding of attitude, remains to be seen.

THE UNFOLDING TECHNIQUE

Coombs (1950, 1952) has presented a scaling method in the spirit of scalogram analysis, but dealing with a different situation. The respondent is asked to rank the items in order of preference — or in order of agreement from *strongly agree* to *strongly disagree*. In the model, each item is viewed as having a fixed position on the latent dimension, while each respondent is likewise located on the dimension. The respondent's order of preference is determined by the relative distance of the items from his location. For example, in the situation depicted in Fig. 5, the rank order reported by respondent I would be *BCAD;* respondent II, *BCDA;* respondent III, *CBDA*. Since direction is ignored, it is as if each individual had folded the scale over, using his location as a hinge. From a set of these rank orders, one can deduce the order of the items on the continuum, and thus unfold the continuum, as it were.

It is also possible to deduce something about the relative magnitudes of intervals between items. Note that for each pair of items, the item preferred by a respondent depends on the loca-

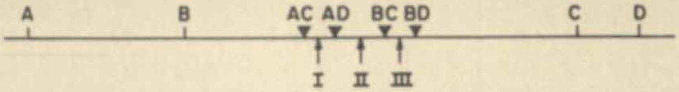


FIG. 5. The model of the unfolding technique. The letters represent items, the Roman numerals represent respondents. The triangles indicate the mid-points of intervals between items.

tion of the respondent in relation to the mid-point of the interval between the items on the scale. Thus, respondent I prefers *A* to *D* because he is on the *A* side of the *AD* mid-point. Thus, the preference orders tell us something about these mid-points. In our case, we know that *AD* is to the left of *BC*. This implies that the interval between *A* and *B* is larger than the interval between *C* and *D*. If respondent II had reported *CBAD,* we would have had to place mid-point *AD* to the right of mid-point *BC*. In this case, the interval between *A* and *B* would have been smaller than the interval between *C* and *D*.

Practical methods of analyzing such a set of data have not yet been worked out. One major difficulty is how to treat errors. As in the case of scalogram analysis, we must first decide which response patterns to accept and which to class as error.

In the case of attitude measurement, it would seem that an additional piece of information would be helpful. The respondent could be asked to indicate which statements were more *pro* than his position, and which were more *con*. This would certainly help in the unfolding process.

LATENT STRUCTURE ANALYSIS

Lazarsfeld (1950a, b) has recently reformulated the theory of attitude measurement. His latent structure theory is a general mathematical model that helps to clarify the problem of attitude measurement. Latent class analysis and the latent distance model are two special cases of the general model that will be discussed in detail.

The general latent structure model. We have defined an attitude operationally as a syndrome of response consistency. Response consistency implies that the manifest items used to assess the attitude are mutually intercorrelated. In fact, the interrelationships among items define the attitude. The attitude variable must get its meaning from these consistencies. In designing a mathematical model for the latent attitude variable, we shall simply reverse this reasoning, and say that the items are intercorrelated *because* of the latent variable. Since each item is related to the same underlying variable, the items must be related to each other. Moreover, this is the only reason they are related to each other. The latent variable accounts for all of the manifest interrelationships. If the latent variable were held constant, or partialled out, the items would be independent. This is the basic idea in latent structure analysis.

Although the latent structure model is applicable to any type of manifest variable, our discussion will be simplified if we consider only dichotomous items. The operating characteristics of these items can be represented in terms of the probability of a positive response. Let x represent the latent attitude variable. Then let $f_i(x)$ be the operating characteristic of item i, as a function of x. (Lazarsfeld's term is "trace line.") For any fixed x_0, $f_i(x_0)$ gives the probability of a positive response to item i. Now let $\phi(x)$ be the distribution of scores on the latent variable. For any fixed interval about x, $(x, x + dx)$, $\phi(x)dx$ gives the proportion of people having latent attitude scores in that interval; i.e., $\phi(x)$ is a probability density function. Then the proportion of people who have attitude scores in an interval about x and who respond positively to item i is $f_i(x)\phi(x)\,dx$. To find the total proportion who respond positively to item i, we must integrate over all possible values of x:

$$p_i = \int_{-\infty}^{+\infty} f_i(x)\phi(x)dx.$$

Now consider the proportion who respond positively to both items i and j. The basic assumption of latent structure analysis is that the items are independent if x is held constant. But $f_i(x)$ and $f_j(x)$ are the probabilities of response for a fixed x. Thus, for any x, the probability of a positive response to both items i and j is the product $f_i(x)\,f_j(x)$. Then the proportion of people who have an attitude score in an interval about x and who respond positively to items i and j, is $f_i(x)f_j(x)\phi(x)dx$. Finally,

$$p_{ij} = \int_{-\infty}^{+\infty} f_i(x)f_j(x)\phi(x)dx.$$

In a similar manner, we can obtain

$$p_{ijk} = \int_{-\infty}^{+\infty} f_i(x)\, f_j(x)\, f_k(x)\phi(x)dx.$$

The complete system of these equations, for p_i, p_{ij}, p_{ijk}, p_{ijkm}, etc., completely describes the item interrelationships.

Very little can be done with the model without making further assumptions about the operating characteristics, or the attitude distribution, or both. By making such assumptions, several special models may be derived. The single common factor models presented in the discussion of Likert's scaling method are examples of special latent structure models.

The similarity of latent structure analysis and factor analysis has been demonstrated by Green (1952). Thurstone's (1947) multiple factor analysis can be derived as a special case of the general latent structure model, extended to continuous manifest variables.

Latent class analysis. One of the important models stemming from the general theory is latent class analysis. The basic notion of this model is that the interrelationships among the manifest items reflect the existence of two or more distinct subpopulations which make up the total population of respondents. Within each subpopulation, or latent class, it is assumed that the items are independent. However, the proportions of people who respond positively to each item will vary from class to class. It is this variation of item marginals which makes for correlation between items.

As an example, consider two items concerning government ownership of industry:

1. Should the government own the public utilities?
2. Should the government control the coal mines?

For some population of respondents we might get the following fourfold table of proportions, indicating a positive correlation between the two items:

		Item 1		
		+	−	*Total*
	+	24	23	47
Item 2	−	17	36	53
	Total	41	59	100

Suppose now that the population is made up of a group of prosocialists and a group of antisocialists. It is not claimed that the classification is clear-cut and abrupt, but rather that two latent classes with these general tendencies can be found. Then the table for the total population might be broken down as follows:

Prosocialists

	+	−	*Total*
+	20	12	32
−	5	3	8
Total	25	15	40

+

Antisocialists

	+	−	*Total*
+	4	11	15
−	12	33	45
Total	16	44	60

=

Total

	+	−	*Total*
+	24	23	47
−	17	36	53
Total	41	59	100

Within each latent class, the items are independent. Knowing the response to one item does not help to predict the response to the other. For the total group, however, there is a positive relationship between the items. The existence of the two latent classes is responsible for this correlation.

Within each class there is still considerable lack of response determination. Presumably there are several other factors operating which lead a certain proportion of the people in the class to give each response. However, since the items are independent, the factors influencing the split on one item must be different from the factors influencing the split on the other item. For, if there were any factor tending to lead to a positive response on both items, this would lead to a correlation between the items for that class.

In the mathematical model, the attitude distribution is represented by the parameters n_s, the proportion of people in class s. The item operating characteristics become the parameters

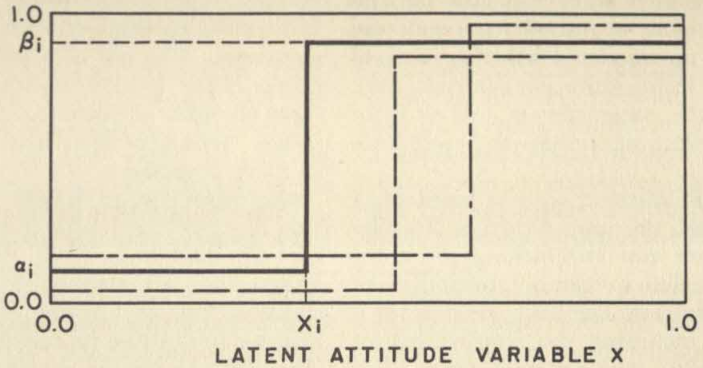


FIG. 6. Representative item operating characteristics in the latent distance scale. The three parameters (α_i, β_i, and x_i) are shown for the solid characteristic in the figure.

v_{is}, the probability of a positive response to item i for class s. We then have directly that

$$p_i = \sum_{s=1}^{m} v_{is} n_s,$$

$$p_{ij} = \sum_{s=1}^{m} v_{is} v_{js} n_s,$$

$$p_{ijk} = \sum_{s=1}^{m} v_{is} v_{js} v_{ks} n_s.$$

Methods for estimating the latent parameters n_s and v_{is} have been developed by Lazarsfeld and Dudman (Chapter 3 in Lazarsfeld, *et al.*, 1950b), Gibson (Chapter 6 in Lazarsfeld, *et al.*, 1950b), and Green (1951a, b). The procedures become quite complicated if more than two or three classes are involved. Anderson (1953) has developed the sampling theory of the estimation procedure devised by Lazarsfeld and Dudman.

The latent class model implies a nominal attitude scale, with each individual being placed in one of several latent classes. Bayes' rule can be used to compute the probability that an individual belongs to a particular latent class, on the basis of his item responses. The classification can be made according to the maximum probability. However, this classification rule may not lead to classes with the appropriate parameters. If a set of data fits the latent class model, then we know that the data behave *as if* there were a number of latent classes with the stated properties. However, we do not know whether the population of respondents can be subdivided into classes having these properties. This logical difficulty is inherent in the model.

Many psychologists prefer to consider attitude as a continuous variable rather than a discrete variable having only a few values. It is, of course, possible to consider the discrete class model as an approximation to the continuous model.

There are several unsolved problems in latent class analysis, including item selection and statistical tests of goodness of fit. Also, there have been very few attempts to use the model. Rossi (1951) and Green (1951b) applied the model to problems of social stratification. Until more empirical evidence has accumulated, it is impossible to evaluate the utility of the class model.

The latent distance scale. The latent structure model that provides a probability model for scalogram analysis is called the latent distance scale. Although the latent distance scale can be written as a special case of the latent class model, its relation to the Guttman scale will be clearer if we treat it as a continuous model.

For convenience, the range of the attitude variable is restricted to the interval $0 \leq x \leq 1$, and the attitude distribution $\phi(x)$ is assumed to be uniform over that interval. That is, $\phi(x) = 1$. Each item operating characteristic is assumed to have some low constant value, α_i, up to some breaking point, x_i, and to have a high constant probability, β_i, thereafter. Examples of such trace lines are given in Fig. 6. A simpler model can be obtained by assuming $\beta_i = 1 - \alpha_i$. Methods for estimating the parameters x_i and α_i have been developed by Lazarsfeld and Hays (1952).

The perfect Guttman scale is obtained by setting α_i equal to zero (and β_i equal to unity). However, the utility of the latent distance model is that an imperfect Guttman scale can be represented equally well. Not only can the error in each item be evaluated relative to the total scale, but the model enables us to compute the scale from summary statistics. We can determine the parameters x_i and α_i from the data in the form of the sets of proportions p_i and p_{ij}.

Since the latent distance scale is based on a probability model, the nonscale types of scalogram analysis are not troublesome — every response pattern has a certain probability of occurring. In addition, the many criteria for a Guttman scale, including the requirement of 90% reproducibility, become unnecessary. For the latent distance scale, we merely wish to know whether the model fits the data within the limits of statistical fluctuation. An arc sin chi-square test similar to that used in the method of paired comparisons would be appropriate.

RATING METHODS

Some investigators attempt to measure attitude by asking a judge to rate the attitude of each individual. In some cases, a single over-all rating is made. Alternatively, the judge rates the individual on several specific items, from which a composite measure is obtained.

The information on which a rating is based may have several sources. Krech and Crutchfield (1948) list seven possible sources: nonverbal behavior toward the object, verbal behavior relating to the object, secondary expressive cues, responses in clinical-type interviews, personal documents, responses on projective tests, and immediate experience. If the ratings are to be made in a uniform manner, it is essential that the same kind of information be available for each individual. The use of a standardized interview would help to ensure uniformity of information. Reliance on haphazard observation is more likely to increase error variance than to lead to an insightful rating.

A single over-all rating can be obtained in several ways. Each individual can be rated on a discrete scale, such as a ten-point scale. In a variant of this technique, the position of an individual is indicated on a continuous line segment that represents the attitude continuum. Verbal descriptions attached to points on the scale help to ensure the consistency of ratings. An over-all rating can be obtained from the respondent himself. Riker (1944) found correlations of about .65 between self-ratings and typical Thurstone scales.

If several ratings can be obtained for each individual, any of the judgment methods can be used to get a scale of individuals. Here the individuals are treated as items to be sorted and compared. The use of a judgment method requires either a large number of judges, or a large number of repeated ratings by a single judge. Neither of these alternatives is practical in many situations.

One objection to single, over-all ratings is that the rater may not be using the same criteria for all individuals. We might find, for example, that *A*'s attitude toward the political status of Negroes was being compared with *B*'s attitude toward the economic status of Negroes. This difficulty can be overcome by dividing the total situation into several specific components, and obtaining a rating of each individual on each component. The components then become items that can be scaled by one of the judgment or response methods. For example, rather than rate on the basis of a general impression obtained from a TAT, one could use specific items, such as: "Does the protocol indicate aggressive tendencies toward Jews?" At a more specific level, one could use items such as: "Was adultery mentioned or implied in any of the stories?" Behavior rating scales are of the same general nature. The use of specific items of this sort gives the experimenter some control over the criteria of the judges, and helps to ensure uniform treatment of all respondents.

The forced-choice technique (Sisson, 1948; Travers, 1951) has recently been proposed as a method for systematic rating. On the basis of item analysis data, pairs of items are chosen which have equal popularity but different discriminating power. Two such pairs are combined in a tetrad. The rater is required to say which item in the tetrad is most like the individual and which is least like the individual. Gordon (1951) gives this example from a personality questionnaire: (a) a good social mixer, (b) gives in readily to other people's wishes, (c) thorough in any work undertaken, (d) worries about possible misfortune. The major difficulty with this technique is that the four items in a tetrad may not be comparable. Of course, if all items refer to the same underlying dimension, this criticism is not valid.

The forced-choice technique is one attempt to overcome the halo effect. It is well known that some sets of specific items are not much better than an over-all rating scale, since all the items can be rated on the basis of a general impression of the individual. By forcing a decision between items that would otherwise both be checked, the effect of halo can be reduced. Whenever possible, however, items should be constructed which are sufficiently objective so that halo becomes a minor factor. In this case, the forced-choice technique would not be particularly useful.

The rating methods introduce a new aspect of reliability, since different judges will undoubtedly give different ratings. We can determine stability for one judge, as well as equivalence and homogeneity, if specific items are used, and we can also determine rater reliability by correlating the ratings of different judges. (In the mental test literature, the problem is called reader reliability; see Gulliksen 1950b.) To determine rater reliability, at least two judges must rate each individual. If rater reliability is found to be high, then in future studies it is reasonable to allow different raters to rate different individuals. Ebel (1951) has presented a method for determining the average rater reliability if several raters are used.

COMPARISON OF SCALING METHODS

Scale criteria. Each scaling method is based on a model which relates the data to a latent attitude continuum. Procedures are incorporated for checking the agreement of the data and the model.

There are two sets of data to be evaluated in the judgment methods. In the first place, the scale of items depends on a model for the judgments about the items. In the method of successive intervals, the judgment distributions must be normal on the derived scale. In the method of paired comparisons, the difference in scale values should agree with the proportion of times one of each pair of items is preferred. Secondly, the criterion of relevance is used to check the agreement of the item responses with the assumptions about the item operating characteristics.

The method of summated ratings requires a factor analysis of the items to determine whether a single common factor accounts for most of the inter-item correlation. Item-test correlations are sometimes used in lieu of a factor analysis. An alternative check is to determine the homogeneity of the items.

Scalogram analysis includes several criteria that a scale must satisfy. The reproducibility must be at least .90, the item marginals must have a large range but must not include extreme values, each response category must have more nonerror than error, and errors must be random.

In latent structure analysis, the items must be independent (uncorrelated) for individuals with the same attitude score. No practical statistical test is available to test this hypothesis.

It is apparent from this summary that most of the scale criteria are subjective. The decision depends heavily on the judgment of the experimenter. It is doubtful that the decision could ever be completely automatic, but more objective criteria should be provided for the experimenter. Since, in theory, all the scaling methods can be subsumed under latent structure analysis, a practical statistical test of the latent structure hypothesis could be used with any method.

Unidimensionality. All the scaling models considered here, with the exception of latent class analysis, are unidimensional models. The data are referred to a single underlying dimension. If the data fit the model, the scale is unidimensional by definition. If the data do not fit the model, then either a scale does not exist, or a multidimensional scale is necessary. Unidimensionality has no meaning apart from the context of a particular scaling model.

One might think that latent structure theory has solved the problem of unidimensionality in general, by requiring the items to be independent for a fixed score on the latent attribute. However, we have seen that the general latent structure theory is completely abstract. Additional assumptions must be made about the operating characteristics or the attitude distribution in order to make use of the model. Whether or not a set of data is unidimensional will depend on what assumptions are made. By assuming sufficiently complex item characteristics, any set of item response data can be referred to a single continuum. Dimensionality cannot be divorced from the particular model.

There is an interesting difference between the judgment methods and the response methods concerning unidimensionality. In the response methods, we must have a unidimensional set of items. The response to each item must be a function of the single latent variable, plus some factor unique to the item. But in the judgment methods, the items can be multidimensional stimuli that are judged with respect to a single dimension. Perfectly consistent judgments of brightness can be obtained for colors which differ in saturation and hue as well as brightness. This implies that in some circumstances we could get two different unidimensional scales from the same set of items, using the judgment methods. This apparent paradox is not a logical inconsistency; in fact, it is a useful property of the judgment methods.

Items. The judgment methods require nonmonotone items, while scalogram analysis requires monotone items. Mosteller (1949) has shown that this is not a fundamental difference. He presented a modified scalogram technique for nonmonotone items, and suggested that it would be equally possible to modify the judgment methods so that they could handle monotone items.

Typically, the judgment methods and the method of summated ratings deal with large numbers of items, while scalogram analysis seldom employs more than 10 or 15 items. The single common factor model implies that a large number of items is desirable with the method of summated ratings, in order that the specific item factors will be greatly outweighed by the common factor. Scalogram analysis permits very little error in any item, and hence appears to need fewer items. This is a happy circumstance, since it is very difficult to find the almost perfect item required for a Guttman scale. The importance of the single factor theory is its demonstration that a large number of mediocre items can be just as good as a small number of precise items.

The scaling methods include techniques for discarding items which, for one reason or another, do not meet the requirements of the scale. This is a perfectly logical procedure, but it poses some problems for the description of the attitude universe. We must not only include all the possible items similar to those retained in the questionnaire, but we must also exclude all items similar to those rejected from the scale. In other words, by rejecting items we may restrict or change the nature of the variable being measured. This consideration does not hold, of course, for items which are rejected simply because they are of the wrong form for some particular scaling method.

Metric. The methods of paired comparisons and successive intervals appear to yield interval scales, as does the method of equal-appearing intervals if we make the assumption that equal-appearing intervals are equal. In a narrow sense, the method of summated ratings can be said to provide an interval scale, although it has interval properties only in terms of the population attitude distribution. The Guttman scale, as well as the latent distance scale, is primarily an ordinal scale, although a metric can be provided by the maximum variance method. Latent classes form a nominal scale.

An interval scale should have the property that the score distribution is invariant except for a linear transformation, with respect to the particular set of items used. It should not be possible to manipulate the score distribution by selecting items with special characteristics. According to this criterion probably none of the scales has equal intervals.

In theory, the judgment scales should be invariant with respect to item parameters. If a biased set of items is used, the number of items endorsed by an individual may be changed markedly, but his scale position should not be affected. Actually, it is doubtful that the invariance could be demonstrated. The scale of items could probably be manipulated by a biased selection of items to be judged. Studies of the anchoring effect, summarized by McGarvey (1943), suggest that the distribution of judgments is influenced by the set of items. However, this is not likely to alter the final scale appreciably, since the scale depends on the *relative* placement of items. That is, in the method of successive intervals, a biased selection of items could result in large shifts in the positions of the category boundaries, but the relative shifts of item scale value would probably be small. Thus we shall not be far amiss in treating the judgment scales as interval scales.

Not so with the method of summated ratings, or the maximum variance scales. It is well known that the distribution of scores on these scales is very sensitive to the parameters of the particular items, such as item-total score correlation and item popularity. Skewed, flat, or normal distributions can be obtained by appropriate item selection. The units of the scale have no meaning apart from the particular set of items, and the associated score distribution.

Although the distribution of scores on a Guttman scale is of little importance in most cases, it is equally susceptible to manipulation. Smith (1950) pointed out that the successive item popularities correspond to the cumulative score distribution for a perfect scale.

Experimental comparisons. A number of studies have compared the various scaling methods. Hevner (1930) compared the methods of rank order, equal-appearing intervals, and paired comparisons. In general, she found very close linear relationships among scales constructed by the three methods. However, she found an end effect in the equal-appearing interval scale, both extremes of the scale being compressed. Saffir (1937) compared the methods of rank order, successive intervals, and paired comparisons. He found linear relationships throughout.

Several authors have attempted to compare the judgment methods and the method of summated ratings. Likert, Roslow, and Murphy (1934) used the five-category response with a Thurstone scale. After eliminating the neutral items, they were able to score the Thurstone scale by the Likert method of summated ratings. They found this score to be more reliable than the Thurstone-type score, although the two scores were highly correlated. The items originally used by Thurstone and Chave in constructing their scale of attitude toward the church were employed by Edwards and Kenney (1946) to compare the Thurstone and Likert techniques. They constructed two equivalent 20-item Thurstone scales, which correlated .88, and one 25-item Likert scale that had a reliability of .94. The Likert scale correlated .74 and .92, respectively, with the two Thurstone scales. Only five items selected for the Likert scale were also selected for the Thurstone scales, although this restriction was not imposed. Eysenck and Crown (1949) constructed a 24-item scale of anti-Semitism by the method of equal-appearing intervals. They scored the scale in the usual Thurstone way, obtaining a reliability of .83. If the scale was scored by the Likert method, the reliability was .90. They also tried a "scale product" score that was obtained by multiplying the item scale value by the response score (1, 2, 3, 4, 5) and summing over all items. This hybrid score had a reliability of .94.

These authors seem to have been unaware of the logically incompatible requirements of the two methods concerning item operating characteristics. Theoretically, a nonmonotone type of item could not be used on a Likert-type scale, and a Likert-type item could not be used on a Thurstone scale. In practice, Thurstone items with scale values near the extremes tend to have monotonic operating characteristics. Thus if "troublesome" items with scale values near the center are rejected, the two methods are compatible in practice. Nevertheless, these studies cannot be regarded as evidence of the comparability of Thurstone and Likert scales. The scale product score is logically inconsistent with the nonmonotone item philosophy.

Clark and Kriedt (1948) compared scalogram analysis with the Likert approach. They used the Rundquist-Sletto 22-item scale of economic conservatism, a Likert-type scale. They found marked unreliability in the manner in which response categories were combined. The reproducibilities for two samples were .79 and .80, although the odd-even reliabilities were .85 and .88, and the median item-total score correlations were .58 and .65. Thus there seems to be evidence of a strong common factor, but the scale does not meet Guttman's requirements. Eysenck and Crown (1949), in their study of an anti-Semitism scale, compared the Likert scoring of the scale with a scalogram analysis. They computed a factor analysis of item tetrachoric correlations, and found a strong common factor, with a very weak second factor. The scalogram analysis, however, resulted in a reproducibility of only .85. Item reproducibilities were highly related to common factor loadings with a curvilinear correlation (eta) of .84. The authors asserted that the common factor rather than the reproducibility is the proper criterion of scalability.

Niven (1953) has compared scalogram analysis with the method of reciprocal averages, the latter yielding an approximation to maximum variance scores. From 29 items he selected 10 for the Guttman scale and 12 for the reciprocal averages scale, using objective criteria. The two selected scales had 9 items in common. He found a correlation of .93 between the two scales in the original sample, and a correlation of .89 in a cross-validation sample. This result is consistent with the mathematical analysis that demonstrates that the maximum variance scores, or first principal component, is a monotonic function of the scalogram scores for a perfect Guttman scale.

CONCLUSION

It is clear that there are several procedures available to the investigator who wishes to measure attitudes. The judgment methods, the method of summated ratings, scalogram analysis, and the rating methods are all useful techniques. At present the latent structure models and the unfolding technique must be classed as scaling models rather than as scaling methods, since satisfactory practical methods of application are not currently available. (It seems likely that straightforward procedures for applying the latent class model and the latent distance model may soon be available.)

There is still a great deal to be learned about the scaling methods. Currently, an investigator who must choose one of the methods can find very little information on which to base his choice. More experimental and analytical comparisons of the methods are needed. Latent structure analysis seems to provide a very promising framework for future development of scaling techniques. Multidimensional scaling methods also seem promising for attitude measurement.

In this forest of methods, the investigator should not lose sight of the fact that a scale is made from items. The initial and basic problem of attitude measurement is to assemble a set of carefully worded, insightful items that cover the area in question. A hodgepodge of carelessly contrived items all thrown into the hopper of any scaling technique will invariably result in hash. To obtain reliable, valid scales we need good scaling techniques, to be sure, but we also need good items.

References

Allport, F. H., & Hanchett, Gertrude A. The war-producing behavior of citizens: a scale of measurement with preliminary results in imagined situations. *J. soc. Psychol.,* 1940, *11,* 447-490.

Allport, G. W. Attitudes. In C. Murchison (Ed.), *A handbook of social psychology.* Worcester: Clark Univer. Press, 1935. Pp. 798-844.

Anderson, T. W. On estimation of parameters in latent structure analysis. *Psychometrika,* 1954, in press.

Attneave, F. A method of graded dichotomies for the scaling of judgments. *Psychol. Rev.,* 1949, *56,* 334-340.

Bennett, J. F. A method for determining the dimensionality of a set of rank orders. Unpublished doctoral dissertation, Univ. of Michigan, 1951.

Bennett, J. F. A method for multidimensional item scaling. *Amer. Psychologist,* 1952, *7,* 409. (Abstract)

Bergmann, G., & Spence, K. W. The logic of psychological measurement. *Psychol. Rev.,* 1944, *51,* 1-24.

Bernberg, R. E. The direction of perception technique of attitude measurement. *Int. J. Opin. & Attitude Res.,* 1951, *5,* 397-407.

Bogardus, E. S. Measuring social distance. *J. appl. Sociol.,* 1925, *9,* 299-308.

Borgatta, E. F., & Hays, D. G. Some limitations on the arbitrary classification of non-scale response patterns in a Guttman scale. *Publ. Opin. Quart.,* 1952, *16,* 410-416.

Bray, D. W. The prediction of behavior from two attitude scales. *J. abnorm. soc. Psychol.,* 1950, *45,* 64-84.

Campbell, D. T. The indirect assessment of social attitudes. *Psychol. Bull.,* 1950, *47,* 15-38.

Cantril, H. The intensity of an attitude. *J. abnorm. soc. Psychol.,* 1946, *41,* 129-135.

Clark, K. E., & Kriedt, P. H. An application of Guttman's new scaling techniques to an attitude questionnaire. *Educ. psychol. Measmt.,* 1948, *8,* 215-223.

Coombs, C. H. Psychological scaling without a unit of measurement. *Psychol. Rev.,* 1950, *57,* 145-158.

Coombs, C. H. A theory of psychological scaling. *Engineering Research Institute,* Bulletin No. 34. Ann Arbor, Mich.: Univ. of Michigan Press, 1952.

Cronbach, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika,* 1951, *16,* 297-334.

Davis, F. B. Item selection techniques. In E. F. Lindquist (Ed.), *Educational measurement.* Washington, D. C.: American Council on Education, 1951. Pp. 266-328.

Davis, F. B. Item analysis in relation to educational and psychological testing. *Psychol. Bull.,* 1952, *49,* 97-119.

Doob, L. W. *Public opinion and propaganda.* New York: Henry Holt, 1948.

Droba, D. D. *A scale for measuring attitude toward war.* Chicago: Univ. of Chicago Press, 1930.

Ebel, R. L. Estimation of the reliability of ratings. *Psychometrika,* 1951, *16,* 407-424.

Edgerton, H. A., & Kolbe, L. E. The method of minimum variation for the combination of criteria. *Psychometrika,* 1936, *1,* 183-187.

Edwards, A. L. A critique of neutral items in attitude scale construction. *Psychol. Rev.,* 1946, *53,* 159-169.

Edwards, A. L. On Guttman's scale analysis. *Educ. psychol. Measmt.,* 1948, *8,* 313-318.

Edwards, A. L. The scaling of stimuli by the method of successive intervals. *J. appl. Psychol.,* 1952, *36,* 118-122.

Edwards, A. L., & Kenney, K. C. A comparison of the Thurstone and Likert techniques of attitude scale construction. *J. appl. Psychol.,* 1946, *30,* 72-83.

Edwards, A. L., & Kilpatrick, F. P. Scale analysis and the measurement of social attitudes. *Psychometrika,* 1948a, *13,* 99-114.

Edwards, A. L., & Kilpatrick, F. P. A technique for the construction of attitude scales. *J. appl. Psychol.,* 1948b, *32,* 374-384.

Edwards, A. L., & Thurstone, L. L. An internal consistency check for the method of successive intervals. *Psychometrika,* 1952, *17,* 169-180.

Eysenck, H. J., & Crown, S. An experimental study in opinion-attitude methodology. *Int. J. Opin. & Attitude Res.,* 1949, *3,* 47-86.

Ferguson, A., Myers, C. S., Bartlett, R. J., *et al.* Quantitative estimation of sensory events. *Final Report, Adv. of Sci.,* 1940, *2,* 331-349.

Ferguson, G. A. The factorial interpretation of test difficulty. *Psychometrika,* 1941, *6,* 323-329.

Ferguson, L. W. The influence of individual attitudes on construction of an attitude scale. *J. soc. Psychol.,* 1935, *6,* 115-117.

Festinger, L. The treatment of qualitative data by "scale analysis." *Psychol. Bull.,* 1947, *44,* 149-161.

Ford, R. N. A rapid scoring procedure for scaling attitude questions. *Publ. Opin. Quart.,* 1950, *14,* 507-532.

Fromme, A. On the use of certain qualitative methods of attitude research: a study of opinions on the methods of preventing war. *J. soc. Psychol.,* 1941, *13,* 429-459.

Fuson, W. M. Attitudes: a note on the concept and its research context. *Amer. sociol. Rev.*, 1942, *7*, 856-857.

Garner, W. R., & Hake, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, *58*, 446-459.

Goodenough, W. H. A technique for scale analysis. *Educ. psychol. Measmt.*, 1944, *4*, 179-190.

Gordon, L. V. Validities of the forced-choice and questionnaire methods of personality measurement. *J. appl. Psychol.*, 1951, *35*, 407-412.

Green, B. F. A general solution for the latent class model of latent structure analysis. *Psychometrika*, 1951a, *16*, 151-166.

Green, B. F. Latent class analysis: a general solution and an empirical evaluation. Unpublished doctor's dissertation, Princeton Univ., 1951b.

Green, B. F. Latent structure analysis and its relation to factor analysis. *J. Amer. statist. Ass.*, 1952, *47*, 71-76.

Green, B. F. A method of scalogram analysis using summary statistics. *Amer. Psychologist*, 1953, *8*, 357. (Abstract)

Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.

Gulliksen, H. Paired comparisons and the logic of measurement. *Psychol. Rev.*, 1946, *53*, 199-213.

Gulliksen, H. Intrinsic validity. *Amer. Psychologist*, 1950a, *5*, 511-517.

Gulliksen, H. *Theory of mental tests.* New York: Wiley, 1950b.

Gulliksen, H. A least-squares solution for successive intervals. *Amer. Psychologist*, 1952, *7*, 408. (Abstract)

Guttman, L. The quantification of a class of attributes: a theory and method of scale construction. In P. Horst *et al.*, *The prediction of personal adjustment.* New York: Bulletin No. 48, Social Science Research Council, 1941. Pp. 319-348.

Guttman, L. A basis for scaling qualitative data. *Amer. sociol. Rev.*, 1944, *9*, 139-150.

Guttman, L. A basis for analyzing test-retest reliability. *Psychometrika*, 1945, *10*, 255-282.

Guttman, L. The test-retest reliability of qualitative data. *Psychometrika*, 1946, *11*, 81-95.

Guttman, L. The Cornell technique for scale and intensity analysis. *Educ. psychol. Measmt.*, 1947, *7*, 247-279.

Guttman, L. The problem of attitude and opinion measurement. In S. A. Stouffer *et al.*, *Measurement and prediction.* Princeton, N. J.: Princeton University Press, 1950a. Pp. 46-59.

Guttman, L. The basis for scalogram analysis. In S. A. Stouffer *et al.*, *Measurement and prediction.* Princeton, N. J.: Princeton University Press, 1950b. Pp. 60-90.

Guttman, L. The principal components of scale analysis In S. A. Stouffer *et al.*, *Measurement and prediction.* Princeton, N. J.: Princeton University Press, 1950c. Pp. 312-361.

Guttman, L. The third component of scalable attitudes. *Int. J. Opin. & Attitude Res.*, 1950d, *4*, 285-287. (Abstract)

Hammond, K. R. Measuring attitude by error-choice: an indirect method. *J. abnorm. soc. Psychol.*, 1948, *43*, 38-48.

Henry, A. F. A method of classifying non-scale response patterns in a Guttman scale. *Publ. Opin. Quart.*, 1952, *16*, 94-106.

Hevner, Kate. An empirical study of three psychophysical methods. *J. gen. Psychol.*, 1930, *4*, 191-212.

Hinckley, E. D. The influence of individual opinion on construction of an attitude scale. *J. soc. Psychol.*, 1932, *3*, 283-296.

Horowitz, R. E., & Horowitz, E.L. The development of social attitudes in children. *Sociometry*, 1938, *1*, 301-338.

Horst, P. Measuring complex attitudes. *J. soc. Psychol.*, 1935, *6*, 369-374.

Horst, P. Obtaining a composite measure from different measures of the same attribute. *Psychometrika*, 1936, *1*, 53-60.

Hovland, C. I., & Sherif, M. Judgmental phenomena and scales of attitude measurement: item displacement in Thurstone scales. *J. abnorm. soc. Psychol.*, 1952, *47*, 822-832.

Johnson, H. M. Pseudomathematics in the mental and social sciences. *Amer. J. Psychol.*, 1936, *48*, 342-351.

Kahn, L. H., & Bodine, A. J. Guttman scale analysis by means of IBM equipment. *Educ. psychol. Measmt.*, 1951, *11*, 288-314.

Katz, D. The measurement of intensity. In H. Cantril (Ed.), *Gauging public opinion.* Princeton, N. J.: Princeton Univer. Press, 1944. Pp. 51-65.

Kelley, E. L., & Fiske, D. W. *The prediction of performance in clinical psychology.* Ann Arbor, Mich.: Univ. of Michigan Press, 1951.

Kelley, H. H., Hovland, C. I., Schwartz, M., & Abelson, R. P. The influence of judges' attitudes in three methods of attitude scaling. Paper read at Eastern Psychol. Ass., Boston, April 25, 1953.

Krech, D., & Crutchfield, R. S. *Theory and problems of social psychology.* New York: McGraw-Hill, 1948.

Kriedt, P. H., & Clark, K. E. "Item analysis" versus "scale analysis." *J. appl. Psychol.*, 1949, *33*, 114-121.

LaPiere, R. T. Attitudes vs. actions. *Soc. Forces*, 1934, *14*, 230-237.

Lazarsfeld, P. F. The logic and mathematical foundation of latent structure analysis. In S. A. Stouffer *et al.*, *Measurement and prediction.* Princeton, N. J.: Princeton Univer. Press, 1950a. Pp. 362-412.

Lazarsfeld, P. F., *et al.* Columbia-RAND project report, 1950b.

Lazarsfeld, P. F., & Hays, D. G. A system of scales derived from a standard trace-line model. Columbia-RAND project report, 1952.

Likert, R. A technique for the measurement of attitudes. *Arch. Psychol.*, 1932, No. 140.

Likert, R., Roslow, S., & Murphy, G. A simple and reliable method of scoring the Thurstone attitude scales. *J. soc. Psychol.*, 1934, *5*, 228-238.

Loevinger, Jane A systematic approach to the construction and evaluation of tests of ability. *Psychol. Monogr.*, 1947, *61*, No. 4.

Loevinger, Jane The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychol. Bull.*, 1948, *45*, 507-529.

Lord, F. M. A theory of test scores. *Psychometric Monograph No. 7.* Richmond, Va.: The William Byrd Press, 1952.

Lord, F. M. The relation of test score to the ability underlying the test. *Educ. psychol. Measmt.*, 1953, *13*, 517-549.

Lorge, I. The Thurstone attitude scales: I. Reliability and consistency of rejection and acceptance. *J. soc. Psychol.*, 1939, *10*, 187-198.

Lorge, I. The fundamental nature of measurement. In E. F. Lindquist (Ed.), *Educational measurement.* Washington, D. C.: American Council on Education, 1951. Pp. 533-559.

MacCorquodale, K., & Meehl, P. E. On a distinction between hypothetical constructs and intervening variables. *Psychol. Rev.*, 1948, *55*, 95-107.

McGarvey, H. R. Anchoring effects in the absolute judgment of verbal materials. *Arch. Psychol.*, 1943, *39*, 1-80.

McNemar, Q. Opinion-attitude methodology. *Psychol. Bull.*, 1946, *43*, 289-374.

Marder, E. Linear segments: a technique for scalogram analysis. *Publ. Opin. Quart.*, 1952, *16*, 417-431.

May, M. A., & Hartshorne, H. First steps toward a scale for measuring attitudes. *J. educ. Psychol.*, 1926, *17*, 145-162.

Mosier, C. I. A modification of the method of successive intervals. *Psychometrika*, 1940a, *5*, 101-107.

Mosier, C. I. Psychophysics and mental test theory. Fundamental postulates and elementary theorems. *Psychol. Rev.*, 1940b, *47*, 355-366.

Mosteller, F. A theory of scalogram analysis using non-cumulative types of items. Report No. 9. Laboratory of Social Relations. Harvard University, 1949.

Mosteller, F. Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations. *Psychometrika*, 1951a, *16*, 3-10.

Mosteller, F. Remarks on the method of paired comparisons: III. A test of significance for paired comparisons when equal correlations and equal standard deviations are assumed. *Psychometrika*, 1951b, *16*, 207-218.

Mussen, P. H. The reliability and validity of the Horowitz Faces test. *J. abnorm. soc. Psychol.*, 1950, *45*, 504-506.

Newcomb, T. M. The influence of attitude climate upon some determinants of information. *J. abnorm. soc. Psychol.*, 1946, *41*, 291-302.

Niven, J. R. A comparison of two attitude scaling techniques. *Educ. psychol. Measmt.*, 1953, *13*, 65-76.

Noland, E. W. Worker attitude and industrial absenteeism: a statistical appraisal. *Amer. sociol. Rev.*, 1945, *10*, 503-510.

OSS Assessment Staff. *The assessment of men.* New York: Rinehart, 1948.

Pace, C. R. A situations test to measure social-political-economic attitudes. *J. soc. Psychol.*, 1939, *10*, 331-344.

Payne, S. L. Thoughts about meaningless questions. *Publ. Opin. Quart.*, 1950, *14*, 687-696.

Payne, S. L. *The art of asking questions.* Princeton, N. J.: Princeton Univer. Press, 1951.

Pintner, R., & Forlano, G. The influence of attitude upon scaling of attitude items. *J. soc. Psychol.*, 1937, *8*, 39-45.

Postman, L., & Zimmerman, C. Intensity of attitude as a determinant of decision time. *Amer. J. Psychol.*, 1945, *58*, 510-518.

Proshansky, H. A projective method for the study of attitude. *J. abnorm. soc. Psychol.*, 1943, *38*, 383-395.

Riker, B. L. Comparison of methods in attitude research. *J. abnorm. soc. Psychol.*, 1944, *39*, 24-42.

Rosander, A. C. An attitude scale based upon behavior situations. *J. soc. Psychol.*, 1937, *8*, 3-15.

Rossi, P. H. Latent structure analysis and research on social stratification. Unpublished doctor's dissertation, Columbia Univer., 1951.

Saffir, M. A. A comparative study of scales constructed by three psychophysical methods. *Psychometrika*, 1937, *2*, 179-198.

Sanford, F. H. The use of a projective device in attitude surveying. *Publ. Opin. Quart.*, 1950, *14*, 697-709.

Sherif, M. An experimental approach to the study of attitudes. *Sociometry*, 1937, *1*, 90-98.

Sisson, E. D. Forced choice—the new army rating. *Personnel Psychol.*, 1948, *1*, 365-381.

Smith, G. H. Beliefs in statements labelled fact and rumor. *J. abnorm. soc. Psychol.*, 1947, *42*, 80-90.

Smith, R. G., Jr. Reproducible scales and the assumption of normality. *Educ. psychol. Measmt.*, 1950, *10*, 395-399.

Snedecor, G. W. *Statistical methods* (4th Ed.). Ames, Iowa: Iowa State College Press, 1946.

Stevens, S. S. Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 1-49.

Stouffer, S. A., *et al. Measurement and prediction.* Princeton, N. J.: Princeton Univer. Press, 1950a.

Stouffer, S. A. An overview of the contributions to scaling and scale theory. In S. A. Stouffer *et al.*, *Measurement and prediction.* Princeton, N. J.: Princeton Univer. Press, 1950b. Pp. 3-45.

Stouffer, S. A., Borgatta, E. F., Hays, D. G., & Henry, A. F. A technique for improving cumulative scales. *Publ. Opin. Quart.*, 1952, *16*, 273-291.

Suchman, E. A. The scalogram board technique. In S. A. Stouffer *et al., Measurement and prediction.* Princeton, N. J.: Princeton Univer. Press, 1950a. Pp. 91-121.

Suchman, E. A. The intensity component in attitude and opinion research. In S. A. Stouffer *et al., Measurement and prediction.* Princeton, N. J.: Princeton Univer. Press, 1950b. Pp. 213-276.

Thistlethwaite, D. Attitude and structure as factors in the distortion of reasoning. *J. abnorm. soc. Psychol.*, 1950, *45*, 442-458.

Thurstone, L. L. Psychophysical analysis. *Amer. J. Psychol.*, 1927, *38*, 368-389.

Thurstone, L. L. The measurement of opinion. *J. abnorm. soc. Psychol.*, 1928a, *22*, 415-430.

Thurstone, L. L. Attitudes can be measured. *Amer. J. Sociol.*, 1928b, *33*, 529-554.

Thurstone, L. L. Rank order as a psychophysical method. *J. exp. Psychol.*, 1931, *14*, 187-201.

Thurstone, L. L. *Multiple factor analysis.* Chicago: Univer. of Chicago Press, 1947.

Thurstone, L. L., & Chave, E. J. *The measurement of attitude.* Chicago: Univer. of Chicago Press, 1929.

Torgerson, W. S. Multidimensional scaling: I. Theory and method. *Psychometrika*, 1952, *17*, 401-419.

Travers, R. M. W. A critical review of the validity and rationale of the forced-choice technique. *Psychol. Bull.*, 1951, *48*, 62-70.

Tucker, L. R. Maximum validity of a test with equivalent items. *Psychometrika*, 1946, *11*, 1-13.

Weschler, I. R., & Bernberg, R. E. Indirect methods of attitude measurement. *Int. J. Opin. & Attitude Res.*, 1950, *4*, 209-218.

Wherry, R. J., & Gaylord, R. H. Factor pattern of test items and tests as a function of the correlation coefficient, content, difficulty, and constant error factors. *Psychometrika*, 1944, *9*, 237-244.

Wilks, S. S. Weighting systems for linear functions of correlated variables when there is no dependent variable. *Psychometrika*, 1938, *3*, 23-40.

CHAPTER 10

Systematic Observational Techniques

Roger W. Heyns and Ronald Lippitt
University of Michigan

The psychologists of Wundt's day were among the first to be concerned with the relationship between the subjective reaction and report of the human organism and the stimulus situation. In the psychophysical experiments which dominated the last century, the physical characteristics of the stimulus were known by independent techniques. The task was to discover their correlates in experience. In these cases, the subject did not serve the function of determining "what is out there."

The student of behavior moved from this concern to an interest in more complex behavior and to aspects of the organism other than sensory function. Paralleling this shift in interest has been a change in the role of at least some of the human organisms in the experimental process. They became the measuring instruments, analogous to those independent techniques which the earlier experimenter used to determine the physical characteristics of the stimulus situation. As the scientist became interested in problems such as social interaction in a playground situation or the way in which a collection of people solved a common problem, he found it necessary to pay attention to variables for which no physical measuring instruments were available. If the development of sympathetic behavior is an important concern, how can sympathetic behavior be recorded and measured? It is in this connection that the human observer emerges as a measuring instrument. Not only does he define the stimulus situation, but he also records the responses.

This use of the human being as a source of information not about himself but about someone or something else is now widespread. There is hardly an experiment on group process or a study of social behavior which does not call heavily upon his services. In addition, the variety of situations in which he is used has increased considerably and now ranges from carefully controlled laboratory experiments to diverse field situations.

While it is by no means a new approach to the problem of the human observer, it is useful to keep in mind the essential identity between the human observer and the thermometer or other instrument for measuring the stimulus situation. If we do this we shall see the problems of calibration, of agreement, and of validity as essentially the same as in the case of other measuring instruments. While there are undoubtedly some unique problems in the case of the human measuring instrument, the objective is the same as with any other instrument, namely, accurate and reliable information.

Although the use of the human observer has become extensive, there is a relative paucity of information concerning factors which affect the accuracy and reliability of observer data. Strictly methodological studies involving the use of the human observer are rare. Many of the methodological problems have been left untouched largely because most of the experimenters who have used the human instrument have been concerned with using him effectively in their particular situation. No one has made it his concern to study the scope and limits of the human observer *per se*. As a result, the present chapter offers tentative conclusions largely based on the experiences of experimenters rather than on the results of precise methodological investigation.

A Brief History of Observer Methods

The informal history of the social observer goes back to the beginning of time. The novelist and the poet have long been concerned with the precise description of social behavior and with deriving propositions about motives and feelings from observed behavior. The average citizen in his day-to-day interactions makes precise observations, especially in areas which are of critical importance to him. The insecure employee makes inferences concerning the feelings of his boss and these inferences are based upon precise descriptions of the way in which the boss opens the door and closes it, his facial expressions, and so on. For the socially effective individual these observations and inferences from them are, by and large, shared. When they become notoriously inaccurate or generally unshared we regard the individual as a deviant. Part of the learning task of every child involves accurately noting and interpreting complex social cues. Observation of children indicates that the very small child learns rather quickly the significance of some very subtle cues. In fact, part of our theorizing about child development, especially in the area of parent-child relationships, is based on the proposition that the child responds in terms of valid inferences he makes concerning the motives and feelings of those around him. If we accept this general argument, the extensive use of the trained observer is primarily an application in the scientific area of a general skill which most humans have to some degree. The application of this ability to the pursuit of knowledge imposes some refinements and some more rigorous requirements, but no new or occult skill.

The modern, or scientific, history of the human observer, as far as the social sciences are concerned, probably begins with such researchers as Bühler, Stern, Dorothy Swaine Thomas, Florence Goodenough, and Olson. The origins of the trained observer of behavior in medicine go back much further, of course. The impetus for quantification of observations was first apparent in the work of Thomas and her associates (Thomas, 1933; Thomas, Loomis, and Arrington, 1933). As a group, these early investigators used the human observer in certain ways which might be contrasted with much of the more recent use. The early emphasis tended to be on exhaustive observational categories. The major focus was on the formal, rather than functional, character of behavior. The categories involved little interpretation on the part of the observers. These experimenters were primarily concerned with physical acts such as pushing, touching. Another central concern, particularly in the case of the Thomas group, was with reliability as measured by agreement between observers. This concern with reliability led to progressive modification of the categories of description until the inter-rater agreement was at a maximum. As a result, the categories of observation became more and more removed from any theoretical basis and their predictive efficiency was minimal. Recent studies using the observer differ from the pioneering efforts as, more and more, the theory underlying the research, rather than reliability considerations, has come to dictate what is to be observed. The movement has been away from exhaustive recording of all the behavior which occurs toward the selection of particular aspects of behavior in order to answer specific questions. This trend has resulted in the development of different observer systems, depending on the purpose of the experimenter and the framework he is using. The newer systems also differ from the earlier ones in that they more frequently require the observer to make inferences concerning motives and feelings from the behavior. We shall discuss in a later section examples of observer systems which exemplify these trends.

OCCASIONS FOR THE USE OF OBSERVERS

We have specified the general condition under which observers have been found to be useful, namely, when other measurement devices are unavailable or inappropriate. We shall turn in this section to a more precise description of the circumstances in which the human measuring instrument is necessary or useful.

When other techniques of measurement would disrupt the process of social interaction. It might be perfectly possible to ask the participants in a group discussion at various points during the meeting how satisfied they are with the behavior of the leader. However, this procedure might disrupt the process under study in two ways. The first would be particularly important when the investigator was concerned with the occurrence or nonoccurrence of certain events later on in the group discussion. The fact that the group is stopped and has to resume may have an effect on the time at which these events occur, and might, in certain circumstances, prevent their occurrence. Here the effect of the interruption is not on the variable under observation but on some other dimension of group process. The second kind of disruption concerns whether or not the measurement itself produces changes in the variable with which the investigator is concerned. The measuring instrument calls the attention of group members to a particular aspect of their behavior or their reactions. This particular characteristic may, as a result, become salient for them, and the knowledge of the experimenter's interest in a particular area may have an influence on the reactions which they later report. This would be particularly true when the measuring instrument calls attention to aspects of behavior which might lack social approval. In these cases the members of the group might inhibit such behavior. Consequently, the interruption for measurement produces a change in the variable being measured.

When the actor himself is inadequate as a direct source of information. At the risk of oversimplification, we may suggest that there are basically only two ways in which we investigate human behavior. We watch the human organism or we ask him questions. There are a variety of ways in which we ask questions, some of them quite direct, as in a self-rating scale, and some of them very indirect, as in a projective test. Sometimes the individual scores himself, as with an attitude scale; or his performance is scored by someone else, as with a projective protocol. There are, however, a variety of circumstances in which asking the actor is unsatisfactory, no matter what the method.

One of the most obvious of these is when the score or measure in which the experimenter is interested cannot be derived from individual scores. An illustration of this type of score is *rate of interaction.* In this case the experimenter might be concerned with the pacing of the meeting and use as indices the extent to which there is a delay between the participation of one member and the next participation, or the frequency with which more than one person is speaking.

Using the actor as an informant may also be inadequate because the phenomenon with which the experimenter is concerned is ephemeral. If it is embedded in a prolonged and complex series of interactions it is likely that the memory of the informant is sorely strained or totally inadequate for the task. It may be a temporally brief event likely to be overlooked by the group member who is intent on other objectives.

The actor as a source of information frequently falls short because he lacks the sensitivity or the vocabulary which the particular observation requires. Most category systems—indeed, most theoretical systems—which are being applied to complex social interaction situations call attention to dimensions which are not the ones in which the typical group member reports social interaction. Typically, when one asks an untrained observer what went on at a meeting, he reports in terms of the content of the discussion. To make the actor useful as a source of information on some of the less obvious dimensions of group process he must be sensitized to the existence of these dimensions.

The circumstances may be such as to make it unlikely that a valid report would be made by the actor. In a study of seventy-two conference groups in business and industry, the vast majority of the participants indicated that they were highly satisfied with the leader's performance. Since in most cases the leader was also the superior in the organization to the members, there is reason to suspect that there were pressures against reporting dissatisfaction with the leader. A variant of this would be the case where the experimenter was concerned with the incidence of behavior which was unconsciously motivated. There are situations in which personality theory would predict that the actor's report of his own behavior would be distorted. Similar distortion might occur in the actor's report of the behavior of other people toward him. There is ample evidence that perception is selective and in part determined by personality needs. When the investigator is interested in behavior which might trigger off anxiety responses in the participants, it is hardly likely that their description of this behavior will correspond to what actually occurred.

THE TWO BASIC OBSERVER SYSTEMS: CATEGORY SETS AND RATING SCALES

THE TWO OBSERVER TASKS

There are essentially two tasks which can be given to the observer. He can be asked to classify the behavior he sees into categories or he can be asked to assign some numerical index to that behavior. Fundamentally, these two tasks are probably the same in the sense that when the observer assigns a number to a behavior, he is putting that behavior in a class of behaviors which have that number. In the ideal rating scale, productive of maximum reliability, there is a clear definition of the dimension of observation *and* a specification of the properties which behavior must have in order to be appropriately assigned a number on the scale. It is as though subcategory definitions and illustrations are erected perpendicular to the scale, defining the characteristics of behaviors to which each scale number should be applied.

When we apply a category system to the same task, we decide the zero point; that is, we decide the conditions which must be met for the act to be classified in that category. If an act does not meet the conditions, we classify it into some other category; in the simplest case this other category includes all behaviors not suitably classified in the critical category. Typically, the system is more complicated than this, providing other categories into which the behavior might conceivably be placed. Usually, these other categories are qualitatively different from each other. They are on different dimensions and have no quantifiable relationship to each other. When, however, they do have a quantifiable relationship to each other, as with a continuous category system described below, we have met the conditions of the ideal rating scale.

Thus, while it may appear that rating and classifying are quite different, they are, in a fundamental sense, the same. Both rating and categorizing require careful specification of the dimension or category. If discriminations within a category are necessary or desirable, and these discriminations should be numerically ordered with respect to each other, the rating scale is one solution; a continuous category system is another. If it is necessary or desirable to discriminate among residual behaviors, a set of rating scale items or a category system with qualitative differences between categories are alternative solutions. When an act can be rated simultaneously on several scales we approximate the condition of a category system which permits multiple coding. This is the case where the categories in the set are not mutually exclusive and hence have no logical necessary qualitative or quantitative relationship to each other.

Carter and his associates made a careful study of the relationship between ratings and category scores (Carter *et al.*, 1951a). After his observers had coded a set of behaviors, they were asked to rate, on a seven-point scale, each of the participants on fifteen personal traits and four group-oriented traits. The ratings were correlated with scores arrived at by combining the categories which might be included in the global trait characteristic. Thus, for example, the rating on authoritarianism was correlated with the total incidence in the following categories: disagrees or is skeptical, argues with others, vigorously argues with others, deflates others, gives bald commands or prohibits (in disagreeable fashion). The correlations of five trait ratings with appropriate combined category scores are all positive and range from .30 to .86. The agreements tend to be highest where the reliabilities of the two types of scores are highest. The data Carter reports support the contention that if the two variables are assessed with high reliability there will probably be a strong relationship between ratings of behavior and appropriate category scores.

In spite of the essential similarity of categorizing and rating tasks for observers and the fact that under certain circumstances the rating job and the categorizing jobs can be identical, there are considerations which have led experimenters to assign one rather than the other of these tasks.

The usual assumption in the case of category systems is that all behaviors classified in a given category have equal weight. Thus, an expression of gratification, a joke, and laughter are each entered into the category of Tension Releasing. They are also assumed to have as much weight as any entry into any other category. However, the experimenter might feel, or his theory might dictate, that while all of these behaviors had the same properties, they do differ in significance. He could take this into account by asking the observers to enter the category more than once for the more significant behaviors. As indicated above, the problem could also be handled by a category system along a single dimension whose categories had prescribed numerical relationship to each other. When observations are necessary along a number of dimensions it is often too time-consuming for the observer to use a continuous category system along each one. The alternative is to ask the observer to assign a numerical weight to each of the behaviors, as in a rating scale.

It sometimes happens that the behavior occurs with such rapidity that it is impossible to note each occurrence. It may be desirable under these conditions to have the observer rate a large unit of interaction along the prescribed dimension, rather than have frequency of occurrence be the score, as would be the case with a category system.

By and large, observer ratings have been used primarily because they are easier to obtain and more aspects of social interaction can be covered by a single observer. This is particularly true when large units of interaction are given a single rating. Observer ratings have, however, all of the general deficiencies of ratings. Most serious are those which result from statistical manipulations involving assumptions which the rating data cannot meet. Rating scale data appear more likely to seduce the experimenter into using statistical indices demanding unwarranted assumptions than are category system data.

Finally, rating scales are more often useful in the early stages of research and theory development in a particular area. It is sometimes easier to obtain better agreement among observers with ratings when the dimension of observation is not precisely defined, and the cues are multiple and not clearly set down, than when category systems are used. This lack of precision is, of course, not a desirable state of affairs, and ultimately the point should be reached where clear definitions are available.

THE DIMENSIONS OF OBSERVER SYSTEMS

There are many differences among existing category systems or sets of rating scale items. They differ in the frame of reference the observer is to use, the aspect of social behavior toward which they are directed, the extent to which they call for special training of observers, and many other respects. This section presents a listing of some of the more important formal properties of observer systems. While the subsequent paragraphs are written with category systems uppermost in mind, virtually all of these methodological dimensions apply as well to rating systems. By a rating system we mean not only the specific items in the scale but definitions of the items and the instructions given to raters.

The dimension of exhaustiveness. In the case of some category systems all the behavior can be classified into one of the defined categories in the set. As we shall see, Bales' (1949) interaction categories make up such a system. His objective was to develop categories capable of enveloping all the behavior in a small face-to-face group. A contrasting system was that used by Jack (1934) in her study of ascendance and submission in a play situation. Her set of categories refers to forms of ascendant behavior; behavior not in these categories was not coded. In a sense, of course, all less-than-exhaustive category systems are really exhaustive systems, since behavior not codable into one of the categories is implicitly in a category labeled "not in the system." The distinction is nevertheless a real one and essentially consists of determining how much of the total observable behavior is to be classified into a positively defined set of categories. Typically, in a less-than-exhaustive system, there is no record of the amount of behavior which occurred which was not in the defined categories. This may be a serious deficiency, since there are occasions when scores are meaningless without an estimate of the total activity which has taken place. For example, in a study of social influence in a camp situation by Polansky, Lippitt, and Redl (1949), it was necessary to correct the influence scores of the individual camper for total amount of activity he engaged in with other campers.

The dimension of inference. Observer systems differ in the amount of inference they require from observers. Let us suppose that an experimenter is concerned with the influence of leader behavior on various aspects of the pattern of verbal participation in a problem-solving group. He may ask his observer to make a notation each time a member talks. This is essentially a two-category system: *talking* and *being silent.* It requires practically no inferences on the part of observers, and requires only that the cues indicating talking and those indicating silence be specified. If this same experimenter were concerned with motivation to participate verbally, he might set up categories such as "eager to talk" and "reluctant to talk." This would require that the observers scrutinize not only the verbal behavior but also the nonverbal behavior for cues suggesting one or the other of these tendencies. The observer deduces from the behavior which he sees the underlying orientation toward participation on the part of the members.

In every case, the theory dictates the level of conceptualization and the kinds of observations which must be made. In some cases, reports of directly observable data are sufficient to test hypotheses. In others, inferences must be made from behavior and the hypotheses tested by inferred data. In the latter case, the question still remains as to who is given the task of inference. The experimenter faces the choice of specifying low-inference categories from which he can later make inferences or asking the observers to make inferences on the spot. In both cases, there is the necessity for defining the properties of behavior which permit classification into the high-inference categories. In other words, if the theorizing which permits the experimenter to proceed from low-inference categories to another set of categories of higher inference is sound, it is also sound enough to specify the kinds of behavior which permit the observer to make certain inferences and not others.

While the kinds of things observers are asked to make inferences about vary enormously, the most common are concerned with motives, intentions, or feeling states of either the actor or the recipient of the action.

The number of aspects of social behavior under scrutiny. It is obvious that in virtually all social situations there are a number of processes operating simultaneously. Thus we can view a group discussion in terms of the number and kinds of influence attempts that are being made. We can choose rather to observe the changes in close interpersonal relationships over time. Or a single category system, in the sense of a set of categories given to a single observer, might be directed at both of these processes. In the Bales category system, which will be described in detail later, there are six such dimensions of group interaction being observed.

Discrete versus continuous categories. The relationships between the categories in any system may be of many different sorts. We are particularly concerned here with whether or not the categories have a quantifiable relationship to each other, such that one category might be said to have less of a specified property than any category above it. For purposes of illustration, we might take the following category system, applied to the social behavior of children: Category 1: Frowns and grimaces. Category 2: Uses words indicating unhappiness. Category 3: Cries bitterly. Here we might argue that Category 3 contained behaviors which showed more extreme evidences of distress and pain than did Category 2 or 1. There remains the real problem of the distance between categories, but the system, to the extent that our assumptions are correct, does have the properties of a continuous relationship between categories. Very few of the category systems which have been used have this continuous property. Most of them consist of discrete categories.

Size of unit. Observer systems differ widely in the definition given for the unit of behavior which is to be classified or rated. The range goes from classifying or rating sentences or parts of sentences to classifying total meetings. In the case of systems directed toward quantifying physical acts, the range is typically from single movements to total behavior sequences.

Certain systems specify what might be called natural units: an act, a sentence. In others, the unit of observation is imposed, as with a time unit. While often the natural unit is a small one, this need not be the case. An agenda item or an activity sequence is an illustration of a natural but macroscopic unit. Typically, the unit to be categorized is defined in part at least by the category system being used. That is, the unit is usually defined as that amount of behavior which is classifiable into a single category.

Range of applicability. Some observer systems were established with the deliberate intention of making them useful in a wide variety of situations. Others were developed with specific purposes dictated by a particular theoretical orientation or research design. There is no necessary reason why systems of the latter sort might not be useful to other experimenters. Very frequently, however, since they were established with a particular situation in mind, the definitions of the categories are not appropriate to another situation. Thus, for example, a category system set up to describe the behavior of a teacher in a school situation might not be immediately appropriate to describing a leader in a group discussion. This may not be, however, because the dimensions of teacher behavior are necessarily inappropriate to describing leader behavior, but rather that the kinds of behavior in the leader situation which should be classified in each of the categories might be different from those in the teacher situation. That is, the categories might be the same but the description of the specific behaviors which belonged in each category would differ.

This distinction is actually an important one. The current state of affairs, so far as observer methodology is concerned, is that new observer systems are created whenever existing systems seem inadequate. They are often regarded as inadequate for the wrong reasons; they are discarded not because the categories in an existing system are inappropriate, but because the precise behaviors which are used to define the categories are not appropriate. The net result is that we are gathering very little data which are commensurate from one experiment to another. Often the experimenter need only change the behavioral referents; however, he usually also changes the categories or the dimensions in terms of which observation takes place.

If we look at the dimension of applicability of a system in these terms it becomes clear that some systems are more generally applicable than others in the sense that they define appropriate behaviors at a level of generality which permits their application in a variety of situations. Others are much more limited in their applicability because the behavioral referents are limited to the setting in which they were developed. When the conceptual meaning of the categories is clear, the behavioral manifestations in other situations often can be specified. Thus the question of the applicability of a category system to a variety of situations depends on the clarity of the conceptualization of the categories.

Illustrative Observer Systems

The principal objective of this section is to present a representative set of techniques which have been used, a sampling of the types of problems which have been attacked with the use of observers, and the settings in which observers have functioned. The classification scheme in terms of which the various systems are presented is admittedly a rough one. The major headings are not mutually exclusive; a given observational procedure might well be classified under several headings. This is a problem, however, which arises no matter what classification scheme is proposed. The main justification for the form decided upon here is that typically the experimenter, in search for a method, thinks first about the kind of phenomenon he is concerned with. This, in a general way, corresponds to the classification scheme proposed in this section. The reader is asked to keep in mind the dimensions along which observer systems differ that were suggested in the preceding section.

Most of the systems described in this section are category systems, in spite of the fact that observer rating schemes are probably more frequently used. Our justification for this emphasis is based on two propositions. First, the kinds of rating scales and their limitations are treated in several areas, notably in the industrial and personality literature (Ghiselli and Brown, 1948; Allport, 1937). Social observation has made no unique contribution to rating methodology. Second, although the basic problems are the same for both kinds of observer tasks, the two approaches have developed their own terminology; as a result, a discussion which alternates from one type of task to the other is unduly complicated.

OBSERVATION OF FACE-TO-FACE GROUP INTERACTION

Much of the current interest in observer methodology springs from theoretical and practical interest in small-group functioning. The largest single category of observer systems is concerned with describing aspects of small-group interaction. Systems of this sort run the gamut from those which attempt to categorize virtually everything that goes on to those which describe particular dimensions of a single individual, the leader, for example. The following section describes five systems which are fairly representative of the work that has been done in this area.

Bales' category system. (1) Purpose and theoretical framework. Bales is interested in utilizing the study of small groups to contribute to a body of theory that will be relevant not only to small face-to-face groups but to larger social systems (Bales, 1949). He deliberately set out to develop a standard set of categories of observation which had general utility. While Bales does not object to specific systems, closely tied to the intention of the investigator and the specific kind of group with which he is working, he feels that a general theory of face-to-face group interaction can be developed more readily by the systematic collection of comparable data on a wide variety of groups. This is facilitated by a general category system. Bales also feels that the widespread use of such a general category system will promote the development of norms on significant dimensions of group process.

To achieve this objective of a set of categories which would have wide applicability over groups with different histories, functions, and composition, Bales tried a variety of approaches. He searched the literature for existing category systems. He utilized free observation situations to note the kinds of categories which emerged spontaneously. He deliberately attempted to observe a wide variety of groups, e.g., teaching seminars, leadership training groups. All in all, about a dozen revisions of the original category system took place. At one time there were as many as 85 categories, and at a later time as few as 5. The final system consists of 12 categories. It represents a compromise between the conflicting pressures of theoretical adequacy on the one hand, and such practical considerations as the number and kinds of distinctions which the observer can make, and ease of analysis and interpretation of the resulting data. The category system, as Bales says, consists essentially of a set of hypotheses about the basic dimensions of interaction in face-to-face groups. As such, the possibility exists that it will be modified because of information about groups which the instrument itself provides.

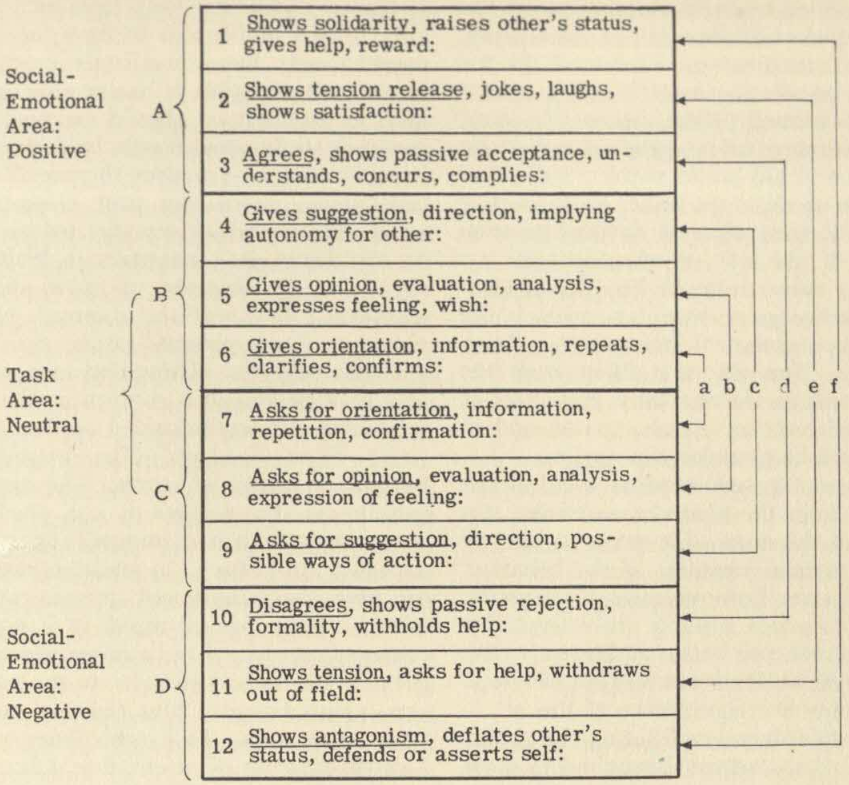
Table 1 shows the categories in Bales' system and the major interrelationships among them. There are six primary dimensions of group interaction under scrutiny: the process of orientation, evaluation, control over others, decision, tension management, and integration. On each of these dimensions, acts can be positive or negative.

Put very briefly, an interacting group faced with a common task faces problems in six areas: problems involved in (1) arriving at a common definition of the situation (orientation); (2) developing a common value system in terms of which alternative solutions are to be evaluated (evaluation); (3) attempts of members to influence each other (control); (4) arriving at a final decision. To meet and solve problems in each of these areas, the group must deal with problems of its own organization. It must (5) deal with the tensions which arise in the group situation (tension management) and (6) work at the task of maintaining an integrated group (integration).

The group process is characterized by oscillation from one of these problems to another. Attempts to deal with difficulties in one of these areas give rise to problems in others. In suc-

TABLE 1

THE SYSTEM OF CATEGORIES USED IN OBSERVATION AND THEIR MAJOR RELATIONS*


Social-Emotional Area: Positive
A
Task Area: Neutral
B
C
Social-Emotional Area: Negative
D
a b c d e f

1 Shows solidarity, raises other's status, gives help, reward:
2 Shows tension release, jokes, laughs, shows satisfaction:
3 Agrees, shows passive acceptance, understands, concurs, complies:
4 Gives suggestion, direction, implying autonomy for other:
5 Gives opinion, evaluation, analysis, expresses feeling, wish:
6 Gives orientation, information, repeats, clarifies, confirms:
7 Asks for orientation, information, repetition, confirmation:
8 Asks for opinion, evaluation, analysis, expression of feeling:
9 Asks for suggestion, direction, possible ways of action:
10 Disagrees, shows passive rejection, formality, withholds help:
11 Shows tension, asks for help, withdraws out of field:
12 Shows antagonism, deflates other's status, defends or asserts self:

KEY:

a Problems of Communication
b Problems of Evaluation
c Problems of Control
d Problems of Decision
e Problems of Tension Reduction
f Problems of Reintegration

A Positive Reactions
B Attempted Answers
C Questions
D Negative Reactions

*From Bales, 1951, p. 59.

cessful groups, the final solution represents a consistent set of solutions to problems in all of these areas.

There is another classification scheme which can also be applied to the categories. The middle categories (4-9) are concerned principally with the problem-solving tasks; the behaviors falling in these categories are emotionally neutral. They are instrumental to performance of the task, whether successful or not. The extreme categories are appropriate for behaviors which are essentially social-emotional in direction; they are concerned with problems of group organization and integration and have emotional loading. Rather than being instrumental to task performance, they are expressive behaviors.

This over-all framework for the category system is important because it means that the ordering of the category is not capricious. To a large extent the definition of a category, and hence the behavior which is to be classified into the category, rests on its position in the entire system. If the total set of categories is not understood, the observer is likely to modify the definition of the categories he does use. This means, practically, that it is as important for the observer to understand and remember the total system and the theoretical relationships among the categories as it is for him to retain in detail the specific definitions for each one.

(2) Orientation of the observer. In this system the observer takes the point of view of the "generalized other." He tries to think of him-

self as a member of the group or, if possible, as a specific member toward whom the action is directed. His job is to classify the behavior of the actor in terms of its significance for the recipient or viewer of the act. "In other words, the observer attempts to put himself in the shoes of the person the actor is acting toward and then asks himself: 'If this fellow (the actor) were acting toward me (a group member) in this way, what would his act mean to me? That is, what is he trying to do, either for himself or for us jointly (i.e., what is the instrumental significance of the act) or what does his act reveal to me about him, or his present emotional or psychological state (i.e., what is the expressive significance of his act)?'" (Bales, 1949, p. 39). The observer differs from the participant only in that he has a special set of categories which direct his observation and in terms of which he gives his descriptions.

The categorizing task requires a minimum of inference from the observer concerning the motivation of the actor. He works at and reacts to the surface meaning of the behavior. The categories are, however, process categories; that is, they are not suitable for categorizing topical or substantive material. Hence it does require that the observer abstract out of a topical contribution the significance of the act in terms of the problems confronting the group. In this sense, the observer is required to move at least one step away from the substantive meaning of the behavior.

(3) Unit of observation. As is typical of most unit designations, the unit of behavior to be classified is defined in this system in part circularly; that is, by the category system itself. "This unit to be scored is the smallest discriminable segment of verbal or nonverbal behavior to which the observer, using the present set of categories after proper training, can assign a classification under conditions of continuous serial scoring" (Bales, 1949, p. 37). The unit is a single interaction capable of appropriate classification into a single category. Typically, a simple sentence is a single unit; complex sentences always contain more than one. Compound sentences are broken up into simple ones and each receives a score.

Nonverbal behavior is also scored, but since it is continuous and not readily broken up into units the procedure calls for assessments of the total group by the observer at one-minute intervals.

In a real sense, then, Bales' system purports to be an exhaustive one. Every act, verbal or nonverbal, is susceptible to classification into one of the 12 categories.

(4) Context. One of the persistent sources of observer disagreement results from differences between observers in the amount of the context of the act which they keep in mind. Bales' answer to this problem is to have the observers consider only the immediately preceding act in their classification of the present one. This kind of decision, it should be noted parenthetically, is dictated in part by decisions made in other areas, as is so often the case. To permit continuous observation and categorizing of small interacts, which was dictated as desirable by the theory, it is necessary to limit the demands on the observer to follow the logical argument. This was also imposed by the decision to have process rather than topical content categories. These decisions, of course, also bear on decisions concerning the amount of inference to be demanded of observers.

(5) Types of scores. The category system provides a variety of scores. The distribution (profile) of the number of acts which fall in each category can be computed for individuals as well as for groups. In addition, ratio scores, for both individuals and groups, can be obtained. Thus, for example, it is possible to compute the ratio of the number of acts in category 7, Asks for orientation, to the number of acts in this category plus those in category 6, Gives orientation. This index Bales suggests as a rough measure of orientation difficulty.

(6) Reliability. Bales has worked intensively on training procedures. He has also developed a manual which defines the categories and makes the system communicable. With intensive training, Bales reports agreement coefficients ranging from .75 to .95.

Conference research problem-solving category system. (1) Purpose and theoretical framework. This category system was developed as part of the methodology for studying the decision-making conference; the specific objective was that of describing the problem-solving process (Heyns, 1948). The categories are defined in terms of intellective functions which must be performed in order for a group to solve a problem and arrive at a decision. The categories and their definitions are as follows:

Goal setting: These contributions have the function of establishing or of suggesting goals or objectives, both of procedure and of content. They are concerned with ends to be attained. These objectives, goals, or ends may be those of the individual, which he is trying to have the group attain; they may consist of statements of accepted goals of the group or part of the group.

Problem proposals: These contributions serve the function of presenting a problem, either in content or procedure. These contributions are concerned with means to ends or goals.

Information seeking: These contributions have the function of seeking to obtain information of an objective, factual, or technical nature. The information sought is from the area of fact on which the group decision is to be based. Contributions seeking factual, objective, or technical information concerning the procedure of the group or an individual are also classified here.

Information giving: These contributions have the function of providing objective, factual, or technical information, either in the subject area or with respect to procedure. The category includes the citing of examples or illustrations.

Solution proposals: These contributions serve the function of indicating solutions to problems. They are suggested means to ends. Modifications of or additions to the solution proposals previously offered are classified in this category if the context gives the contribution a solution proposing function.

Development seeking: These contributions serve the function of attempting to obtain clarification of previous contributions. They seek to determine what was intended by a previous contribution, what its implications are, what inferences are permissible. These frequently take the form of an inference stated as a question. Also included here are contributions which facilitate the procedure of the group by asking the group as a whole (or individuals) to comment, indications to individuals that they have the floor, etc.

Development giving: Contributions here elaborate, make explicit, enlarge on contributions. Included here are: inferences from previous contributions, self-repetitions or restatements by others of previous contributions; reflecting types of contributions which are distillations of previous contributions without attempt at clarification. Finally, this category includes contributions which provide the rationale, reasons, or arguments for the individual's position. They give his reasons for saying what he does.

Opposing: These contributions are characterized by an opposition, resistance to, or disagreement with a suggestion, solution, interpretation, etc. Responses which point out obstacles, difficulties, or objections are included here.

Supporting: These contributions serve the function of indicating agreement or approval of a suggestion or solution proposal. Included here are indications of approval of the fact that another has contributed, whether approval of content is present or not. This is a supporting comment in procedure.

Summarizing seeking: These contributions ask, in effect, for a summary, e.g., "I'm lost, where are we now?"

Summarizing giving: These contributions summarize the group, or part of the group's progress to date. They refer either to substantive material discussed over a period of time, conclusions reached, or to the group's procedure over a period of time. Summary statements of individual participations are not included here.

Nonproblem directed: This includes statements which have no reference to the subject matter of the conference or to the group procedure.

The assumption underlying the system is that there are, in fact, two kinds of problems which confront any problem-solving group. The first is that of the substantive problem, the agenda item. The second is that of organizing and mobilizing the group to deal with the substantive matter. The system makes the further assumption that the intellective functions which must be performed are the same for both kinds of problems.

(2) Orientations of the observer. The observer is instructed to code the contributions of participants in terms of the theoretical function of the contribution. That is, the observer is told not to make judgments about the intent of the actor, nor to appraise the effect of the contribution on the subsequent process. The observer is outside the process and views each contribution in terms of its theoretical properties as a problem-solving function. Thus, a particular contribution is categorized as solution proposal because it is theoretically a solution to a problem or a means to a goal, not because it necessarily functions as such later in the meeting. The categories involve no inference on the part of the observer concerning motives of the actors or feeling states of recipients of acts. Only verbal behavior is classified.

(3) Unit of observation. As with Bales, the unit of behavior to be categorized is defined in terms of the category system itself. The precise

statement of the unit is "any subject-predicate unit which is classifiable into a single category." This entity is referred to as a functional unit. There may be more than one functional unit in a single participation of a group member.

(4) Context. The same procedure is used as with the Bales system. The observer classifies on the basis of the immediate context of the behavior. Thus, while a particular contribution might be classified as opposing a solution proposal made earlier in the meeting, if in the immediate context of the contribution it has the function of supporting the solution proposal just made it is so classified.

(5) Types of scores. The basic data for a given group meeting consist of a tabulation of the number of contributions falling into each category for each participant. These scores may be analyzed on the level of the individual or they may be combined to yield an over-all picture of the group. Both levels of analysis have been found to be valuable.

For describing the group as a whole, the most obvious score is the percent of the total contributions falling into each category. On the individual level, two scores are possible: scores which describe the individual's performance in relation to self, and those which describe his performance in relation to that of the total group. Hence, for each group member there are two scores: the "% own" score which expresses the concentration of verbal behavior in each category as a fraction of the participant's own total behavior, and the "% category" score which indicates what part of the total contributions of the group in each category is accounted for by a given group member. These two scores reflect behavior patterns which may be quite different conceptually. If a participant accounted for most of the opposing contributions during the meeting, his % category score on opposing would be correspondingly high, but his % own score for opposing need not be the highest of all his % own scores. He may actually stress supporting more than opposing in his verbal behavior, so that the former category has a higher % own score for that individual.

It is readily apparent that both % category and % own scores may be used in designating the roles played by the various group members. The above-mentioned individual who contributed most of the opposing done in the group could be described, relative to the other members, as the "critic." However, in terms of the comparative incidence of functions performed by that individual, he may also be referred to as a "supporter." Two different sets of hypotheses may thus become necessary for dealing with the different role designations arrived at through the use of the two scores.

(6) Reliability. In several studies in which it was possible to determine the extent of agreement between trained observers, the correlation coefficients ranged from .64 to .97, for both types of individual scores. Rank order coefficients for group scores in a series of studies ranged from .71 to .98.

Steinzor's social interaction category system. (1) Purpose and theoretical framework. Steinzor's purpose, like Bales', was to develop a method for studying the basic processes in face-to-face group interaction (Steinzor, 1949). His more specific objective was to develop a system for describing group atmosphere. He starts his pursuit of such a system by noting that behavior is motivated and hence cannot be understood without taking into account motivational factors. Verbal behavior is a highly important aspect of adult social relations and motivational components are present in verbal behavior. Steinzor believes that our understanding of group functioning will be greatest when the inferences that are made concerning the motives of actors in a social situation most closely approximate the way in which the individual himself would define his purpose. In addition, he attempts to develop a system that will be applicable to a wide variety of verbal interaction situations and one whose categories are as free as possible from value connotations. His final system consists of two category sets. The major one is comprised of 18 intent categories and one residual category:

Activate and originate
Structure and delimit
Diagnose by labeling
Evaluate
Analyze and explore
Seek information to learn
Express and give information
Clarify conclusion
Defend
Offer solution
Conciliate
Understand and reflect
Give support
Seek support
Oppose and attack
Show deference
Conform
Entertain
Miscellaneous

The labels listed above roughly describe the categories, but each intent category is accompanied by a description. Thus, for example, the giving of factual data is classified in the "Intent to Structure" category. The second category system describes the direction of the intent. Three loci of intent are specified: self, group, or the issue. Thus, for example, one can ask questions of oneself, ask for personal information from other group members, or ask about group procedures, and, finally, one can ask for information about the issue. This second system can be conceived as a subcategory system within each of the larger ones. Steinzor regards these categories as points on a continuum of self-involvement.

It should be noted that Steinzor's system does not purport to describe the total motivational picture in the group at a given moment. It focuses on the intent of the actor, ignoring the motivational consequences of the act in the other members of the group. By the same token, any group score represents only the motivational characteristics of the actors.

(2) Orientation of the observer. By definition, the observer using this system is required to make inferences concerning the intent or motivation of the actor. This judgment about intent determines the category into which the behavior is to be placed, not the effect of the behavior on subsequent group processes. In the use which has been made of the system to date, observers have worked exclusively from typescripts. Thus the inferences had to be based on the actual content of the remarks and not on nonverbal cues.

(3) Unit of observation. Steinzor defined the unit to be categorized as the entire statement a person made that occurred between the statements of individuals immediately preceding and following the person's expression. If the individual stopped talking for five or more seconds and then continued, without anyone else participating, his statement is counted as two units. In the published research using this system, the observer may assign more than one code to longer statements, but this is left to his discretion.

(4) Context. Steinzor, in his description of his system, does not deal explicitly with the question of how much of the total interaction is to be taken into account in defining the intent of a particular verbal act. He indicates (Steinzor, 1949, p. 119) that the judgment of intent is more difficult the more restricted the behavior on which the judgment is based. This suggests that the observers are to use as much of what has already transpired as they can retain.

(5) Types of scores. As with the two category systems just described, it is possible to obtain both individual and group scores as well as ratio scores.

(6) Reliability. The reliability figures which Steinzor reports are quite low. There are a number of factors which probably combined to produce a low degree of agreement, such as relatively brief training of coders, poor motivation of coders, the loose definition of the unit to be categorized. Finally, the judgment of intent was made on the basis of transcribed material. This leaves out many of the nonverbal cues of motivation on which we depend in everyday life. The presence of these interferences makes it unwarranted to say that the low reliabilities indicate the impossibility of inferring intent. Indeed, there is evidence that a well-trained coder using this system is consistent with himself when the coding job is resumed after a period of two months. The greater success of the category system described below, which requires much the same skill from observers, also indicates that inferences of this sort can be made reliably. At the present time, the over-all evidence concerning the reliability of scores indicates that the Steinzor system requires considerable training of observers and probably some combining or redefinition of categories.

Self-oriented need observational system. (1) Purpose and theoretical framework. This observational system was developed as part of the methodological arsenal used to observe decision-making conferences in business and industry (Fouriezos, Hutt, and Guetzkow, 1950). In contrast to Bales', the focus of this procedure is on a particular aspect of group interaction, namely, the motivation of the participants. According to the developers of this system, the behavior of individuals in a group may be regarded as being generated by either or both of two sources. The behavior may be dictated by the requirements of the group situation or primarily by the individual's own personality needs. The latter is labeled *self-oriented need behavior* and is directed principally to the satisfaction of the need itself, regardless of the effect the behavior has on the attainment of the group goal. The experimenters recognized that both sets of determinants may be operating simultaneously; the principal distinction is that of the degree to which the behavior is a function of one or the other.

Five need categories were selected which according to pilot studies provided the greatest amount of agreement between observers. These are: Dependency, Status, Dominance, Aggression, and Catharsis. The categories are defined fairly precisely by the authors and behavioral cues are suggested for each. The main function of the categories in the system, however, is to provide a framework for the observer. The final score is not the number of entries in each category but rather a rating made by the observers. During the group session itself, the observer tallies particular instances of the expression of the types of needs. At the conclusion of the meeting, he makes a subjective appraisal of his tallies in the form of a single, over-all rating on an eleven-point scale. The scale runs from 0 (nonexpression of self-oriented need) to 10 (all behavior of the self-oriented type).

There is no clear-cut evidence that the device of classification during the meeting, prior to the final rating, is essential to the method. Its principal value is probably that of keeping the observer alerted to the dimensions.

(2) Orientation of the observer. As indicated above, the category system provides a "cue manual" for the observers. The observer takes into account the total situation in which the behavior takes place. He is concerned not only with the content of the contribution but its affect loading. This involves sensitivity to the nuances, the inflections, and timing of comments. The system does not assume that the individuals showing self-oriented need behavior will do so in all situations; as a result there is no pressure on the observer to make a hard and fast personality assessment of each individual which ascribes to him some enduring properties. It is interesting to note, however, that in one study the ratings of an observer who saw the participants only in the group situation correlated .82 with those of observers who saw the individual in the group situation and had had four hours of personal interaction with the participants and knew the results of a clinical appraisal which included Rorschach scores, TAT scores, sentence completion test results, and the results of an intensive interview. This clinical battery was aimed at assessing the same needs as were specified in the category system.

It is clear that this system requires a great deal of inference on the part of observers. Each participation is scrutinized in terms of the possible motivational forces which are operating. The whole process requires making deductions from the behavior of the individual as to what the motivational factors are.

(3) Unit of observation. This is the single participation, in the sense that each participation is categorized. The final score for each person is based upon the over-all impact of the individual's performance and hence is based on the entire meeting.

(4) Types of scores. Again it is possible to derive both individual and group scores. The researchers have used two procedures to arrive at a group score. The first of these consists of an average of all the individual ratings. The second is an assessment of the extent to which the group process as a whole was characterized by self-oriented need behavior. The two group scores correlate .67.

(5) Reliability. In two studies of the degree of agreement between observers in the ratings given to individual participants, the correlations reported are .67 and .73. In another, the rank orders of two groups of 16 participants rated by two observers correlated .86 and .96. The average intercorrelation of three observers of the total group score on a sample of 30 groups was .81. The indications are that this procedure can provide reliable scores with a relatively small amount of training.

Carter's observational procedures. (1) Purpose and theoretical framework. Carter and his associates have been working particularly on the problem of leadership in small face-to-face groups (Carter, Haythorn, Meirowitz, and Lanzetta, 1951a, b). This orientation has dictated the dimensions of group process which are observed. The procedures and the categories of observation were developed in settings where small groups worked on three different tasks in a laboratory situation. One of these tasks was that of assembling objects from available materials. As a result of this kind of observational situation, the Carter system includes many more categories that are applicable to nonverbal behavior than do the systems already described.

Table 2 lists the categories in one of the systems which Carter and his associates have developed. The system contains seven principal dimensions such as showing personal feeling, proposing and initiating action, leader roles in carrying out action, etc., with subcategories under each.

By and large, all the categories are low-inference categories in the sense that the observer need not make judgments concerning the motives or feelings of the individuals. This is less true of the first 12 categories than the others, but even here the observer relies on commonly accepted meanings of behavior and words rather than upon subtle cues.

As was mentioned earlier, Carter and his group have also explored the use of ratings. At the end of the observation session, after coding the behavior in the categories described above, each observer rated each individual on 15 personal traits and 4 group-oriented traits. These ratings were made on a seven-point scale and included such traits as authoritarianism, leadership, initiative, submissiveness, and insight.

(2) Orientation of the observer. Carter's observers work behind one-way screens. They

TABLE 2

LIST OF CATEGORIES: CARTER

Major = *M*
1–19. *Shows a personal feeling of*
1. Aggressiveness or anger
2. Anxiety or insecurity
3. Attention or readiness
4. Confusion
5. Cooperativeness
6. Deference
7. Dissatisfaction
8. Formality or reserve
9. Friendliness
10. Negativism or rebelliousness
11. Satisfaction or accomplishment
12. Status (superior)
20–39. *Proposes and initiates action*
21. Calls for attention
22. Asks for information or facts
23. Diagnoses situation — makes interpretation
24. Asks for expression of feeling or opinion
25. Proposes course of action for self
26. Proposes course of action for others
27. Supports or gives information regarding his proposal
28. Defends self (or his proposals) from attack
29. Initiates *action* toward problem solving which is continued or followed
30. Supports proposal of another
31. Agrees or approves
32. Gives information
33. Gets insight
34. General discussion *re* task
35. Expression of opinion
40–49. *Disagrees and argues — with a somewhat negative connotation*
40. Disagrees or skeptical
41. Argues with others
42. Vigorously argues with others
43. Deflates others
44. Gives bald commands or prohibits (in disagreeable fashion)

Minor = *L*
50–59. *Leader roles in carrying out action*
50. Gives information on how to carry out action
51. Praises, commends, rewards
52. Expresses a desire that something be done
53. Asks for assistance for others
54. Asks for assistance for self
55. Integrates group behavior
60–69. *Follower and "worker" roles in carrying out action*
60. Follows suggestion or directions
61. Offers to help or helps
62. Imitates others
63. Asks for permission
64. Collaborates with others
65. Answers questions
66. Performance of simple work unit (group oriented)
67. Performance of simple work unit — an independent effort
68. Passively helps
70–79. *Abortive or nonproductive behavior, re problem*
70. Initiates action which is not followed or continued
71. Verbal interplay without outcome
72. Listens, but doesn't express self or enter in
90–99. *Miscellaneous*
90. Stands around doing nothing
91. Engages in "out of field" activity
92. Engages in incidental conversation while working

note the originator of the action, the type of behavior, and, if appropriate, the recipient of the behavior. Their task is to note what the individual is doing, looked at from the standpoint of the outsider and in terms of the function of that behavior in the total situation. That is, the observer does not concern himself with the intent of the actor and typically not with the effect of the act.

(3) Unit of observation. This has been a particular problem to this research group because of the presence in their category system of nonverbal categories and nonaction categories. For example, category 72 is "Listens but doesn't express self or enter in." How frequently should this category be checked in ten minutes of listening? An early attempt to deal with this problem was to call for the assessment of each subject every half-minute on expressions of feeling or behavior which involved no change. The interaction proved too rapid for this solution to be practical. The resolution finally decided upon was to categorize at least one unit of behavior for every change in behavior. Accompanying this decision was that of defining a change in behavior quite narrowly.

Thus, for example, each repetition of a routine task is regarded as a change in behavior. The research group using this procedure feels that the net result of this unit designation is that virtually all action and changes in action are recorded, but that there is a less adequate picture with respect to feelings and inactions.

(4) Context. While Carter does not discuss this explicitly, it is probable that the acts are coded in the context of immediately preceding acts and in terms of the immediate task situation, rather than using any long-range body of impressions which the observer might accumulate.

(5) Reliability. Carter reports an average correlation between two observers' (postmeeting) ratings for all traits as .68. For individual traits the correlations run from zero to .94. The highest reliability coefficients were obtained on the "leadership" ratings and those for "orientation towards group acceptance" were lowest. These were .84 and .34 respectively. These coefficients are corrected for double length since, in most calculations, the actual score used is the average of the two observer ratings.

Carter and his associates compute the reliability of category scores by correlating the number of units of behavior for particular categories awarded each individual subject by independent observers over a designated time period. The resulting figures are also corrected for double length because of the way the final score is computed. The reliability coefficients reported are those of combined category scores, where a number of categories have been combined to provide a particular score. For example, Categories 25, 26, and 29 are combined to provide an initiative score. Coefficients for five such combined category scores range from .20 to .89. Carter's data indicate that the reliability figures vary depending on what task was being performed when the observations were made. This is principally a function of the incidence of behavior in the category involved. Thus, for example, the insight category score has low or indeterminate reliability in the group discussion situation because of the low incidence of this sort of behavior in the task. There is little variation in reliability coefficients when the data are computed separately for 4-man and 8-man groups. In general, the evidence indicates that the system provides reliable scores in the hands of trained observers.

OBSERVATION IN A TWO-PERSON SITUATION

In reality, the use of observers in a two-person situation presents no new problems; in fact, the two-person situation is often less complex and permits greater exploitation of the potentialities of the observer as a measuring device. However, some of the newer observational procedures have been developed in a two-person setting and we shall discuss two of these in this section.

Chapple's procedure for measurement of interaction. (1) Purpose and theoretical framework. The total theory, together with its implications for formal and informal organizations, which underlies the Chapple procedure is too extensive to be discussed here. Chapple is interested in objective description of interaction, and was dissatisfied with interview material as a source of information about the dynamics of interpersonal relationships. As a result of a variety of experiences, Chapple developed the notion that there is a relationship between such objective events as the amount and duration of contacts between individuals on the one hand and emotional relationships between people and their attitudes toward one another on the other (Chapple, 1940).

Proceeding from this point, Chapple and his associates have developed an observation procedure which is concerned exclusively with the measurement of the time aspects of interpersonal relationships. The observer's task is to press a key when an action begins, record how long it continues, and when it terminates. This kind of record is kept for both persons in the situation. The observer makes no inferences about emotional states or motives. Nor does he attempt to code the content of what is said or to describe the act. Chapple deals quite explicitly with these decisions. He believes that feeling states and content categories are actually described by his quantitative measures of the time characteristics of the interaction of two persons. While the following quote does not do justice to the support Chapple marshalls for his position, it does give the reader the flavor of his position:

> If the reader sharpens his powers of observation, he will see that in many cases people whom he does not like or cannot get along with *say* exactly the same things that people he does like say. So actors frequently take a short play, play it first as a tragedy, and then, using the same words, play it as a comedy. Here the language is seen as unimportant, and the timing is the factor which makes the difference in its effect on the audience. (Chapple, 1940, p. 33)

Actually, the time record provides a great deal more information about the interaction process than our brief description of what is recorded would suggest. It is possible to note not only how much each person talked or acted, but also characteristics of their interaction: whether they interrupted each other, whether there were or were not delays between the action of one and the action of the other, and so on. These various indices have been conceptualized to a certain extent (Chapple, 1949). Chapple has described and labeled such scores as the following: tempo (how often a person starts to act); activity or energy (how much longer he talks or responds than is silent and unresponsive); adjustment (the length of his interruptions against his failures to respond); initiative (the frequency with which one person takes the initiative as against the other); dominance (the frequency with which one person outtalks or out-acts the other when there has been an interruption); synchronization (the frequency with which one person interrupts or fails to respond). In addition to these, there are indices of flexibility, ability to listen, and others which are arithmetic combinations of the ones listed above.

Chapple's entire procedure is a marked departure from the general trend toward using the observer to make delicate judgments and to code behavior into highly conceptualized categories. Those who have been in this main stream find it hard to believe that an essentially contentless observational system, concerned with the mechanics of interaction, can have much predictive efficiency. As Chapple himself suggests, whether his system of measurement is complete enough to describe interaction successfully will depend on what it is able to predict and how well it does that prediction job (Chapple, 1949, p. 296). On the basis of work completed, Chapple reports that various of his indices discriminate between effective and ineffective sales personnel, between occupants of different job groups in a company, and between normal and psychiatric patients (Chapple and Lindemann, 1942). On the basis of evidence available, Chapple's results indicate that it would be unwise to move too far in the methodology of observation from the more formal properties of interaction. In a sense, Chapple's work constitutes a return, with more sophistication and more theory, to some of the earlier work with observers.

(2) Unit of observation. It is apparent from the description already given of the procedure that the observer records more than verbal acts. He notes the occurrence of physical acts such as gestures and nods of the head. Chapple deals more explicitly with the definition of the unit than do many developers of observer systems. In his system the unit of observation is the unit of action. He defines a unit of action as that period of activity recorded from the initial change in muscular state until a second change brings the activity to an end (Chapple, 1940, p. 24). There are undoubtedly instances in which this definition brings small comfort to an observer faced with deciding whether an action is a new one or part of an old one. However, in the situation in which the procedure has been most extensively used, namely, the interview, where actions are primarily verbal and movement possibilities are limited, problems of confusion about the beginning and end of an act are undoubtedly minimal.

(3) Reliability. Data bearing on the reliability of the observer record are limited. Chapple indicates that on a half-dozen records done by two observers on two tape recorders, there were "no differences in the means obtained or the number of units recorded" (Chapple, 1949, p. 297).

Description of interaction in a therapy situation. (1) Purpose and theoretical framework. The basic procedures of observation, developed in other settings, are currently being employed in the description of the therapeutic process (Keet, 1948; Bergman, 1951; Osburn, 1951; Dittman, 1952). Most of the work of this sort has been done by analysis of typescripts or of recordings rather than with direct observers. However, this limitation to typescript or record analysis seems to us to have been a result of characteristics of the situation in which therapeutic treatment takes place and not a function of the dimensions of interaction which have been measured. Since there is good reason to think such considerations as the need for privacy can be dealt with satisfactorily and still permit direct observation, it seems useful to discuss one of the techniques that has been developed for describing an aspect of the therapy process.

The main impetus toward description of this special kind of social situation has come from Rogers and his associates. One of the pioneering efforts in this group was made by Snyder, and while the Rogers group and others (Bergman, Dittman) have elaborated and refined the procedures, the essential characteristics of this sort of application are well illustrated in this study (Snyder, 1945).

The theoretical approach which dictated the category system which Snyder developed is that which underlies client-centered therapy (Rogers, 1942). The system finally arrived at actually consists of two category sets: one developed for describing the responses of the counselor and the other for describing the client. The counselor category system contains 16 content categories and a residual category for unclassifiable statements. Within this total set are major groupings of categories. An abbreviated statement of the counselor categories, indicating the major groupings, follows:

- Lead-taking categories
  - Structuring
  - Forcing client to choose and develop topic
  - Directive questions; specific type of questions
  - Nondirective leads and questions
- Nondirective response-to-feeling categories
  - Simple acceptance
  - Restatement of content or problem
  - Clarification or recognition of feeling
- Semidirective response-to-feeling category
  - Interpretation
- Directive-counseling categories
  - Approval and encouragement
  - Giving information or explanation
  - Proposing client activity
  - Persuasion
  - Disapproval and criticism
- Minor categories
  - Ending of contact
  - Ending of series
  - Friendly discussion
  - Unclassifiable

Analyses of client responses indicated that many of them had an important second dimension. Some clear content statements were also accompanied by explicit or implicit feeling or emotion. Accordingly, the final set for classifying client responses includes a group of categories having to do with the feeling tone or attitudes of the client. The resulting set, also abbreviated, consists of the following categories:

- Problem category
  - Any statement of problems of symptoms
- Simple response categories
  - Asking for advice or information
  - Answer to a question
  - Simple acceptance
  - Rejection of a clarification or rejection
- Understanding or action-taking categories
  - Understanding or insight
  - Discussion of plans and decisions
- Minor categories
  - Ending of contact
  - Ending of series
  - Not related to problem
  - Friendly discussion
  - Unclassifiable
- Client feeling categories
  - Positive attitude toward self
  - Positive attitude toward counselor or counseling situation
  - Positive attitude toward other persons
  - Negative attitude toward self
  - Negative attitude toward counselor or counseling situation
  - Negative attitude toward other persons
  - Ambivalent attitude toward self
  - Ambivalent attitude toward counselor or counseling situation
  - Ambivalent attitude toward other persons

In the study in which this system was developed and applied, 48 interviews were analyzed; altogether slightly under 10,000 responses were classified. The unit of classification was an "idea," defined as a clearly indicated change in subject matter or attitude.

With the exception of the feeling categories, and these only when the feeling is implicit rather than explicitly stated, the categories require very little inference. Their definitions are clearly stated and illustrated. Thorough study of them, and the framework from which they come, makes the categories quite clearly understood.

No attempt was made in the original study (Snyder, 1945) to determine the relationship between either client or counselor behavior and a success criterion. Snyder's data did provide rich descriptive material on the process itself, the relationship between counselor responses and client responses, and changes in the incidence of certain content categories over time. The findings make good sense and indicate that the system is a very useful one.

(2) Reliability. Data on reliability comes from two sources. Snyder himself repeated the classification of the responses in four interviews more than a month after the original classification. Eighty percent of the items were classified into the identical categories; eighty-five percent were classified in the same major group. Total scores in each category correlated between .76 and .87.

In the second reliability check, a person untrained in psychotherapeutic interviews was given the category definitions. He studied these and the classifications which had been made of the responses in two sample interviews. His scores correlated from .52 to .78 with those of Snyder. These results indicate that the system can be quite readily communicated; with a minimum amount of training moderate reliability was attained. Undoubtedly more training would raise this figure considerably.

OBSERVATION OF SOCIAL BEHAVIOR IN A FIELD SETTING

Up to this point we have been discussing observational procedures developed and applied to small group or two-person interaction under more or less controlled conditions. The subjects are localized geographically and the activities in which they engage are typically limited and prescribed. When the social scientist's interests take him into less controlled situations permitting a greater variety of behaviors, new problems of observations arise. Among other things, the observer must be mobile, and this has implications for the kinds of equipment he can use. Problems of unit definition become more acute, as do problems of identification of actors. The definitions of the dimensions of observation must be clear enough either to make them applicable in a wide variety of activity situations or to indicate their nonapplicability under specified conditions. For example, the investigator may be interested in the influence process. The problem then becomes that of conceptualizing the influence process in such a way as to make influence categories useful and appropriate whether the actors are deciding to go swimming or actually playing ball. A brief review of some of the observation systems which have been developed for use in field situations will illustrate these and other problems.

Study of behavioral contagion in a camp setting. (1) Purpose and theoretical framework. Polansky, Lippitt, and Redl (1949, 1950) were interested in the spread of mood, behavior, or attitude from one person to another through an entire group. This is essentially a problem of influence, and necessitated a procedure for describing the influence process. They define social influence as that phenomenon which occurs when a change in behavior of one individual appears to produce a change in the behavior of a second individual. The first person in this chain is referred to as the actor and the second as the recipient. The investigators were particularly interested in *contagion,* by which they mean the spread of behavior or mood that does not result from direct influence attempts. An instance of contagion is defined as follows: an event in which a recipient's behavior has changed to become more like that of the actor or initiator, and when the change occurs in a social interaction situation in which the actor has not communicated intent to evoke such a change. In the principal study reported, the relationship between group status and influence was investigated.

The observer team in this study used the following categories:

1. Direct attempts at influencing (indicating actor and recipient)
   a. Manner, including: Directs, Demands, Uses force or threat of force, Requests-suggests, Pleads
   b. Goal, including: Personal (actor's own) goal, Group (or shared) goal, Recipient's own goal, Some third person's goal.
   c. Response of the Recipient, including: Accepts and complies, Accepts in modified form, Acquiesces but does not comply, Ignores, Rejects, Could not be determined.
2. Status indicators
   a. Asks permission
   b. Implies superior knowledge or skill in recipient
   c. Ridicules
3. Incidents of contagion
   a. Initiator
   b. Recipients (in order)
   c. Content of the contagion

(2) Orientation of the observer. The campers were divided into groups and, except for periods in which reliability checks were made, one observer followed each group, usually for an entire day. He made his observations through all the activities in which it was a distinguishable group. Three observers were used and were rotated through the groups. Observations were continuous for intervals of 15 minutes. By and large, the categories require a minimum of inference.

(3) The unit of observation. This is not clearly specified as to boundaries, but it is an interaction, especially in the case of direct influence attempts. Typically, in these cases, the unit is a verbal interaction. This particular category system illustrates more clearly than do some others the interdependence between category definition and unit designation. In the case of behavior entering the contagion categories, any act of a group member is potentially a contagion behavior. It becomes one in actuality, however, only when there is a change in the behavior of another person in the direction of the actor's behavior.

(4) Reliability. Reliability checks indicated the need for grouping some of the categories that required differentiations the observers were unable to make in the actual situation. Thus, the broad category "manner of influence at-

tempts" was reduced from five subcategories to two: "directive," which included all the instances in the first three categories; and "nondirective," which included the last two categories of the former system. This is a particularly happy instance of combining categories into broader ones in which the resulting broad categories retain some conceptual clarity. The "response-of-recipient" categories were also reduced from five to two: "successful," including "accepts" and "complies," and "accepts in modified form;" and "unsuccessful," including "acquiesces but does not comply," "ignores," and "rejects."

When the rank order of both actors and recipients of one observer is compared with that of another observer, rank order correlation coefficients rank from .00 to .87, depending on the particular score. Twenty-three of the thirty-four correlations reported are above .60. In general, actor scores are more reliable than recipient scores. This is partly due to less difficulty in identification of actors, especially since in many of the situations of contagion there are many recipients. These data also reflect the fact that two observers may differ in the number of interactions observed. A separate reliability check, when identical interactions were observed, had very high percent agreement figures in categorization, indicating that the final category system itself can be used quite reliably. The difference between the correlations of observer scores assigned to actual children and the percent agreement scores, where the reliability of the observer system itself is under test, illustrates that there can be a real difference between the adequacy with which categorizing is done and the reliability of the scores actually assigned individuals. In this case, the difference is in favor of the system itself, which is contrary to usual experience. It is due in this case to differences between observers in the *extent* to which the system is actually applied to people and not to differences between them in the way they apply it.

As was mentioned earlier, it was necessary to correct individual scores both for total activity and for the amount of time each person was in the presence of others.

Observation methods in psychological ecology. (1) Purpose and theoretical framework. Wright and Barker and their associates have embarked on an extensive investigation, the objective of which is to describe as completely as possible the conditions of life and the behavior of children in a particular small, midwestern town (Wright and Barker, 1950). The procedures of the total study and the theoretical framework guiding this work are too complex to be covered in detail in a chapter such as this. Their study raises some interesting problems and in many ways is a marked departure from the prevailing usage of observers, and for these reasons is included in this survey of existing observer systems.

In the sense that we have been using the term *system*, the observer in this study does not have a system. His job is to describe the behavior of the child being observed and the situation in which the behavior takes place. A brief, quite typical sample of a single record of the behavior of a seven-year-old boy named Raymond and of what happened to him is as follows:

> Stewart said pleasantly, "It will somersault if you let it go now." Stewart seemed to sense Raymond's discontent and sought to placate him.
>
> Raymond agreed immediately that it probably would but listlessly let it go anyway.
>
> He watched closely as the car rolled over and over and stopped at the bottom.
>
> Raymond absently scanned the landscape.
>
> Noticing a large crate in another corner of the basement, he proposed, "Let's go over and play with that crate."
>
> Stewart said flatly, "No."
>
> (The wobbly old crate was constructed of several loosely connected boards and measured about five feet by thirty by thirty inches. It was open completely at one end and the other sides had openings where the boards crossed. The openings were large enough for the children to climb through. The crate was resting on its open end.)
>
> Raymond hopped down from the bank into the excavation and walked purposefully over to the crate.
>
> Strenuously he pulled on the crate. He evidently intended to drag it to the place where the three boys had been playing together.
>
> The crate caught; he had to yank strenuously to move it at all.
>
> He looked at me and said, "I'm a big strong man." He grinned at me as he said this, making a joke of his ineffectiveness.
>
> Raymond absently chanted "Dum, Dum, Dum" as he tugged energetically on the crate.
>
> It stuck again on something in the weeds.
>
> He jerked quickly, leaning all his weight backward as he tried to pull the crate.
>
> At the crate came loose suddenly, he stumbled back a few steps.
>
> He looked at me just briefly.
>
> Raymond grunted as he tugged it with both hands.
>
> (Barker and Wright, 1951, pp. 344-345)

(2) Orientation of the observer. It is apparent from the record above that no special observer skills are particularly demanded by the procedure; certainly none which are outside the repertoire of adults capable of getting along effectively in social situations. One of the objectives, indeed, is to keep theorizing about behavior to a minimum and out of the record. Occasionally, inferences about intentions are recorded but these do not spring from a profound analysis of behavior. The record provides a running description of what a person is doing, what the situation is, and, by inference, what the situation, or parts of it, means to the actor. It also contains, as the example above indicates, relatively minor interpretations and descriptions of the way in which the action was accomplished.

There are several other characteristics of the record which should be noted. There is a description of the situation in which the behavior takes place and aspects of the physical environment—the crate in the sample above. A single sentence covers only one action. In addition, every sentence which describes a particular action is placed alone in a paragraph. If more than one sentence is used to describe the same action, these are grouped into a paragraph. Sentences referring to different actions are never placed in the same paragraph.

The purpose of these specifications concerning the written form of the record is to permit later designation of larger and smaller units of activity. The specified form of the record imposes no limitation on this later task. ". . . its organization will not prejudice any attempt to discover where different units and sequences of action started and stopped, where interruptions occurred, where actions overlapped. . . ." (Barker and Wright, 1951, p. 10).

The unit of observation, at least as it is represented on the form, is a molar act. The context of each act is recorded rather than assumed as in the other systems we have discussed. The observer records as completely as possible every action together with the setting in which it occurs. In the current procedures, to ensure completeness and reduce ambiguity in the record, each observer dictates his observations into a wire recorder in the presence of another member of the research staff. This staff member asks questions about omissions or incomplete reports. A report, edited by the observer, is derived from the original report and the results of the interrogations. This report is, in turn, reviewed by a second staff member for eliminating duplications, and correcting meager accounts. A very detailed final report emerges from this process.

(3) Subsequent analysis. It is beyond the scope of this chapter to describe in detail the kinds of analyses to which these observer data are put. It is clear from the above description that in this procedure the human observer provides the raw data record. Translation of these observations into concepts which have meaning in terms of behavioral theory is done by others. The analysts working on the project have a complex and long set of categories for describing the situation and the behaviors. The amount of agreement between analysts in their descriptions of situations and events in theoretical terms and the extent to which tests of hypotheses are possible on the basis of their coding, do bear on our general topic of controlled observation. The results available are preliminary but indicate that fruitful conceptual descriptions are possible, with a high degree of agreement between coders. This conclusion suggests that often it may be highly desirable to work in this direction of total observation, with no categories of observation at all or only minimally conceptualized ones, leaving the coding into more abstract categories for later. This is particularly true when sophisticated observers are unavailable. It should be clear, however, that the success of the Barker-Wright procedure is no mandate for free and uncontrolled observation. Their observers are very clear about what they are to record and how to record it.

Description of social behavior in terms of interpersonal mechanisms. (1) Purpose and theoretical framework. Coffey and his associates (Coffey, Freedman, Leary and Ossorio, 1950) have developed a category system which has wide applicability. It has been used by observers on the spot and by coders working from typescript (Freedman, Leary, Ossorio, and Coffey, 1951). The system differs markedly from several we have discussed, particularly because the resulting description of social behavior is in terms of interpersonal mechanisms. This is in contrast, for example, to the problem-solving system which yields descriptions of groups in terms of the intellective functions performed.

The system itself developed out of a particular approach to the problem of personality description. This research group conceives of the total personality as consisting of three levels: the public, the conscious, and the private. The levels are defined by the sources of data which describe them. The public level of personality data consists of ratings or categorizations of an individual's behavior and its effects on others. These must be made by objective observers and are independent of the individual's own statements. The conscious level of person-

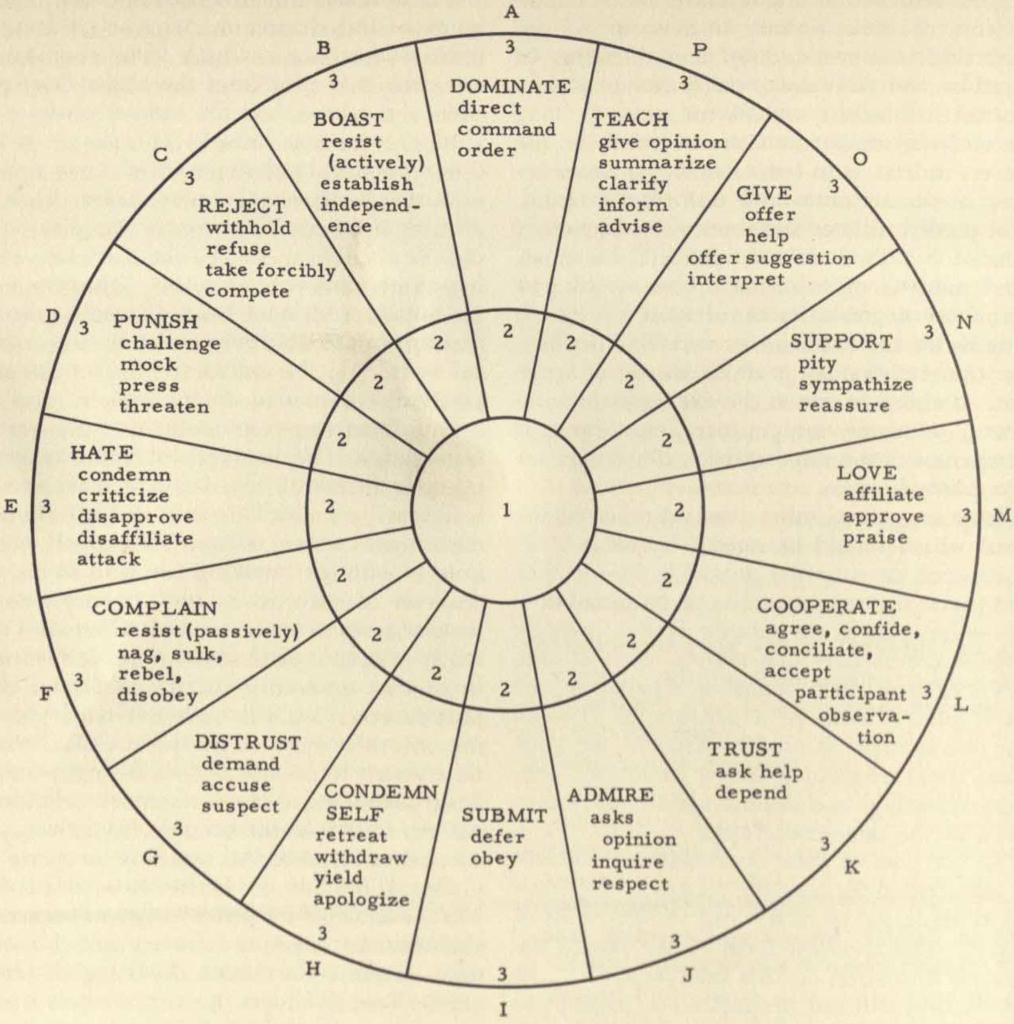


FIG. 1. Level I variables: interpersonal mechanisms with illustrative verbs. Intensity ratings: 3, intense or extreme; 2, average or appropriate; 1, mild intensity. (From Freedman, *et al.*, 1951, p. 151.)

ality data consists of what the individual says about himself and others, at a level of conscious awareness. He may "say" things about himself by means of ratings as well as in unstructured situations. Finally, the private level of data comes from assessment of projective materials such as dreams and TAT stories.

The system described here is that which is required for description of personality at the public level. "The chief concept employed for this purpose is the *interpersonal mechanism*, which is defined as the interpersonal function of a unit of social behavior. The rater or observer asks the following question: What is the subject of the activity, i.e., the individual whose behavior is being rated, doing to the object or objects of the activity? The answer, e.g., he is aggressing against them, or him, affiliating with them, teaching him, etc., is the interpersonal mechanism." (Freedman, *et al.*, 1951, p. 149)

(2) Orientation of the observer. The frame of reference prescribed for the observer in this system is exactly that which the Bales system requires, "the role of the generalized other." The interpersonal mechanisms which comprise the final category system were arrived at after extensive analyses of records. More than eight thousand units of verbal behavior were studied. The system is shown in Fig. 1.

The fact that the categories are presented in the form of a circle is a significant part of the conceptual system. There are four nodal points: dominance, affiliation, submission, and hostility (reading clockwise from the top). Each category falls at the nodal point or is a combination of these nodal mechanisms. Thus, supporting is conceptualized as a blending of affiliation and dominance; rejecting is a combination of hostility and domination.

In addition to categorizing the behavior, the observer is also required to make an intensity rating of the act on a three-point scale. A rating of three indicates an extreme interaction, a rating of two an interaction of average or moderate intensity, and a rating of one is given to interactions of low or mild intensity.

(3) Reliability. The reported reliability data do not permit evaluation of the degree of agreement, category by category. In terms of total scores in all categories, however, the degree of agreement is moderately high. When each of three trained judges' scores was compared with the consensus of the other two, the percentage of complete agreement ranged from 68 to 77.

OBSERVATION OF SOCIAL BEHAVIOR IN CHILDREN

As we indicated in the brief history of systematic observation, most of the pioneers in the use of observers worked with children. Subsequently a large number of observation systems have been developed applicable to the social behavior of children. Only three of the many will be discussed here; the principal justification for the selection probably lies in their representative character and their historical importance.

Measurement of ascendance-submission in children. (1) Purpose and theoretical framework. This procedure was developed by L. M. Jack in connection with an attempt to modify the social behavior of preschool children, particularly along the dimensions of ascendance-submission (Jack, 1934). To determine whether a particular training technique was effective in modifying the incidence of ascendant behavior, it was necessary to develop measures of ascendance. After considerable preliminary work, eight categories of ascendant behavior were specified:

1. Verbal attempt to secure play materials.
2. Forceful attempts to secure play materials.
3. Success in securing play materials.
4. Defending, snatching back materials taken from his possession.
5. Verbal attempts to direct behavior of companion.
6. Companion complies with direction.
7. Forbidding, criticizing, reproving companion.
8. Providing patterns of behavior which companion imitates.

These observational categories were used in an experimental situation where each child was paired with another child. The two children were together in a sand box which contained three sets of toys.

By and large these categories are of the low-inference sort. The observer is called upon to note the occurrence of fairly clear-cut acts, described in terms of everyday language. The significant difference between this category system and earlier ones which also concerned themselves with such low-inference categories is that in this case the behavioral categories are conceptualized. Jack's definition of ascendant behavior is represented in the eight categories.

One other comment about the observer task is necessary. The observers using this system are selective in the observations. There are many behaviors in such a situation which fall outside the category system, e.g., singing to self, playing quietly with toys, striking toy, and so on. The observer is sensitized by the category system to look at certain behaviors and to ignore others.

(2) The unit of observation. We mentioned in another connection that the definition of the unit of observation is usually circular; that is, the unit to be recorded is typically that behavior unit which is classifiable into a single category. This phenomenon of the category definition itself defining the unit is especially clear in such a system as Jack's. An attempt to secure play materials is a unit, no matter how long it takes. The same sequence which involves both verbal and forceful attempts is two units. Finally, if the attempt is successful the same interaction period actually enters the system three times.

We mentioned earlier also that an experimenter often faces the choice of asking for low-inference observations, leaving conceptualization until later, or he can ask his observers to do the conceptualizing. In either case, before the total job is done satisfactorily, there must be a clear specification of behaviors which are to be conceptualized in a particular way. Jack's system illustrates this point nicely. Her category system can be looked at as a set of cues which another experimenter might have used to instruct his observers as to what he meant by ascendant behavior. The behavior descriptions specify cues which indicate that he should make an entry in the category "Ascendant Behavior." In case ratings were used, there would be a subjective weighting of these behaviors to arrive at a final rating.

(3) Reliability and validity. The degree of agreement between observers on total scores assigned to children is quite high. Jack reports

a corrected Spearman Brown correlation coefficient of .80.

Jack reports three relationships which bear on the validity of the ascendance score. First of all, the total ascendance scores based on observation correlated .81 with ascendance ratings made by teachers on a six-item rating scale. The rating scale items were in more general terms than were the behavior categories. Secondly, the observation scores obtained in the experimental situation showed a strong relationship to observational scores on a different set of categories in the playground situation. This second category system was oriented toward assessment of instances of insecurity and lack of confidence. It included such items as showing fear of companion's displeasure, appealing to adult for help, appealing to adult for attention, showing fear of physical objects, and so on. These items appeared four times more frequently in the nonascendant group than they did in the ascendant group. Finally, the ascendance score reflected the effects of a training program which was calculated to increase the amount of ascendant behavior.

Domination and integration in children. (1) Purpose and theoretical framework. Anderson (1939) has done extensive work in the observation of the behavior of children and teachers. His research and the category system are concerned with two concepts: domination and integration. Dominative behavior he defines as a technique of response to others in which the individual resists differences, changes, and growth. The dominative person, according to this scheme, is rigid and inflexible, has his mind made up. He disregards the desires of others, uses commands, threats, or force to gain his objectives. Dynamically, this is the opposite of growth and is the expression of fear of impending change. Integrative behavior, on the other hand, represents maximal growth. Classified as integrative behavior would be: responding voluntarily and without coercion to differences in others, finding common purposes among the differences. The person characterized by social integration achieves a change in structure or function, in goals or purposes as a result of interacting with persons who are different from himself. One of the research problems Anderson set for himself was that of determining the relationship between teacher-child contacts along this dimension of domination-integration, and the nature of child-child interaction along the same dimension. This required the development of an observational instrument for measuring dominative and integrative behavior in teachers' behavior toward children. A similar instrument was developed for observing the interaction of children.

In the case of the teacher-child observation procedures, the final system consists of 20 categories; eight of the categories are classified as domination, nine as integrative, two as ambiguous, and one category for undetermined behaviors. Each contact between the child and the teacher is classified into one of the twenty categories. The observation schedule for use in assessing children's responses to each other consisted of 12 categories, nine of which are classified as dominative and three as integrative behavior. Since seven of the domination categories are identical to those which Jack classified as ascendant behaviors, Table 3 contains both types of categories for the teacher-child instrument but only the integrative categories for the child-child set.

(2) Reliability. Anderson reports a high degree of agreement between observers in the gross number of contacts in both systems, indicating that the problems of specifying the limits of a given contact are minimal. In the case of both the teacher and the child-child system, there is considerable agreement in the classification of dominative behavior. The domination scores assigned by two observers correlate in the .80's. This is also true for the integration score for children. There is less agreement between observers in the case of the integration score for teachers. Anderson reports correlation coefficients of about .50. He attributes this lower degree of agreement to the fact that judgments about integrative contacts are more influenced by their context and hence less specific and more difficult to record. These reliabilities are also reduced in part by the fact that there was a relatively low incidence of entries into these categories.

Anderson, in the monograph which describes his procedure, reports a large number of relationships between integration and domination scores and other variables. His data indicate that the domination scores of two interacting children are highly correlated; domination begets domination. The same is true for integration scores. These and other findings which make sense theoretically indicate that the procedure is useful in the study of social behavior in children.

Measurement of child behavior in terms of social stimulus value. (1) Purpose and theoretical framework. Olson is one of the pioneers in the use of systematic observation in the study of children. This particular system developed out of an attempt to conceptualize personality in terms of the stimulus value of the individual

TABLE 3

ANDERSON'S CATEGORIES OF DOMINATION AND INTEGRATION BEHAVIOR

Teacher System

Domination

- Determines a detail of activity or acts for the child in carrying out a detail
- Direct refusal
- Relocating, repeating, or placing children in different relations to each other or to property
- Postponing, slowing up the child
- Disapproval, blame, or obstruction
- Warning, threats, or conditional promises
- Calls to attention or group activity
- Rations material
- Lecture method (defining a problem or anticipating a question)
- Questions: lecture method (one answer questions)
- Perfunctory questions as statements (indifference)

Integration

- Approval
- Accepts difference
- Extends invitations to activity
- Question or statement regarding child's expressed interest or activity
- Build up (helps child to better definition or solution without giving final answer)
- Participates in joint activity with children
- Gives sympathy
- Gives permission

Children System

Integrative

- Shows common purpose by word or action (includes attempts to cooperate and attempts to share or participate)
- Verbal request or suggestion to direct companions' behavior or secure materials
- Complies with request or suggestion
- Sets pattern including gestures which companion imitates

(Olson and Wilkinson, 1932). The emphasis is that an individual's personality is not completely described in terms of the responses he makes to others but also by the responses others make to him as a stimulus. This approach led to the proposition that important aspects of individual *A* could be derived from recording the behavior of individual *B* toward *A*.

In the specific study reported here, observations of behavior were made in a classroom situation. The ultimate objective was to arrive at scores for the children in the room. What was observed, however, was the behavior of the children and the teacher. The unit of behavior to be recorded was: "such marked restlessness or inattention that the teacher feels the need of speaking to the child or recalling his attention by direct word, look, or gesture." This specification required that the observer single out specific teacher behaviors of the sort described in the definition and record only under certain stimulus conditions, i.e., certain behaviors on the part of the child. During the course of the year 200 observations were made, at various times throughout the day. The total score for each child consisted of the number of occasions, in the 200 observation periods, in which he stimulated the teacher to respond in the way described above.

The system requires a minimum of sophistication from observers, at least as far as the conceptual definition of the behavior to be recorded is concerned. There is little inference required. Since the observer records what is essentially an interaction of a specific sort, there is strong emphasis on the context of the act of the teacher. The significant context is the antecedent behavior of the child. The unit to be recorded is clearly specified.

(2) Reliability. As a result of the low degree of inference and the clear specification of the unit, there is no problem in getting agreement between observers. Total contact scores obtained by two observers correlate .98.

ILLUSTRATIVE EQUIPMENT DEVICES

Along with the efforts to develop observational systems, there has been considerable experimentation with apparatus to aid in one or another aspect of the total observational procedure. In general, equipment devices have been used to deal with three types of problems: improving the efficiency of recording on-the-spot observations, collecting data during the course of the social interaction, and preserving a record of the behavior for later analysis. The principal difference between the last two is that the objective of the devices aimed at data collection during the interaction is typically that of dispensing with the observer. In the other cases, the main purpose is to increase observer efficiency, extend his areas of observation, and promote as much recall of the situation as possible. This is admittedly a rough classification, and some of the devices which have been used on one problem can be used to deal with another; particularly is this true of devices which are here classified in the last category.

DEVICES FOR RECORDING ON-THE-SPOT OBSERVATIONS

There have been a large number of different solutions to the problem of helping the observer record his observations, ranging from simple data sheets listing the category labels or rating scale items to machines which provide the data sheet and a consecutive time record of the interaction process.

Bales' and Gerbrands' interaction recorder. This recorder was developed for use with the Bales category system. It consists of a case containing a machine which drives a wide paper tape (Bales and Gerbrands, 1948). The list of categories fits on the top of the case. On the right side of the list, the moving tape is exposed. When a score is put down, the tape moves under the checklist and disappears, leaving the tape free for another scoring notation. There is a marker inside the case which indicates the end of each minute. In addition, a counter indicates the total elapsed time. The panel of the case also contains a ruby light which flashes on at one-minute intervals. This is the signal for the observer to observe all the members for expressive tension behavior which might otherwise be missed in the observer's concern with more overt interaction. Mechanical details and directions for construction have been published by Bales (Bales, 1948).

Bales and his associates have also experimented with dictating the category and the who-to-whom designation into a sound record which is synchronized with the record of the group's verbal interaction. This procedure is feasible, but it requires a soundproof room for the observers and it sometimes requires the observer to speak while he is listening. There is the additional disadvantage that an hour of recording requires an additional hour of tabulation time.

Stenotype recording. Carter and his associates have worked extensively on techniques for improving the efficiency of recording on-the-spot observations. This group started with the proposition that any form which requires the observer to write necessitates looking away from the group and hence creates the possibility of missing important interaction. This is especially a problem with the Carter system since it requires categorization of nonverbal behavior. One of the devices they have used to meet this and other difficulties has been the Stenograph or Stenotype machine (Carter, 1951b). This is a small manual printing machine used by court stenographers and others for taking dictation. Its keyboard permits the printing, either singly or in combination, of the numbers from 0 to 9 and a number of the letters of the alphabet. As a key is depressed, a two-and-one-half-inch paper tape moves forward one line. This basic unit has been modified to print a time line automatically every half-minute by depressing one of the keys. There have also been minor modifications of the keyboard.

The basic unit and the modifications are relatively inexpensive and the recording tape is very cheap. The tape can be used immediately for tabulation purposes. For optimum use the observer must have memorized the category system and must have mastered the keyboard. The unit can then be used by "touch," as with a typewriter, and the observer can watch the group continuously. The machine can, of course, accommodate many other category systems besides the one with which it has been used. The unit is small, quiet, and easily portable. Observers who have used the system report that it easier than techniques which require dictation.

IBM cards. Lippitt and Zander (1943) developed procedures whereby the observers recorded their observations directly onto IBM cards. In the extensive investigation in which this was done, there were different cards for each area of observation: one for leader behavior, another for total group behavior, and so on. Each observer assigned to a particular area was equipped with copies of the appropriate card. These were placed in a notebook along with code sheets which explained the abbreviations mimeographed on the card itself. Each abbreviation stood for a type of behavior to be recorded. The observer recorded the occurrence of the behavior with a special graphite pencil by drawing a line through the number in parentheses which was opposite the abbreviation on the card.

This procedure has advantages, particularly in the tabulation and analysis phase of research operations. The cards can be fed into a "mark-sensing" machine to produce punched cards which can be used for analysis by other IBM equipment. These advantages, however, accrue only when the data collection job is a large one.

Chapple's Interaction Chronograph. Chapple and his associates have developed an elaborate machine for recording the data required by his system (Chapple, 1949). The machine provides a time record of interaction of two persons as it is observed by the operator. It also acts as a computer, providing summary data on the principal variables in the system. The present machine has 10 columns, each produced by a printing counter, plus a signal counter. From the material fed into and summated by the machine it is possible to get complete scores for individuals separately, as well as scores which characterize their interaction over time. The Interaction Chronograph is not easily moved. Chapple has developed a portable recorder which provides a paper tape record which can be utilized by the computer. In contrast to some of the other devices which have been described, the Interaction Chronograph is designed specifically to provide the data required by the Chapple system and does not have general applicability as a recording device. It is, of course, a specific application of the basic chronograph recording device which does have general applicability.

DEVICES FOR COLLECTING DATA DURING SOCIAL INTERACTION

Strictly speaking, the kinds of equipment to be discussed here are not observational, since the human observer does not provide the data. They are instruments for obtaining data from the actors themselves during the social interaction process.

While there are many variations, with different recording devices being utilized, the basic characteristic of this type of equipment is that the group member is given an opportunity, by pushing a button or pulling a lever, to register his reaction to some aspect of the social scene. Thelen and his associates, who have done the most thorough and extensive work in this general area, have developed a device to record the feeling states of group members which illustrates this general type of mechanism (Thelen and Whitehall, 1949; Thelen, 1950). The group members are trained to push a multiposition lever to indicate their feelings during the meetings. The buttons are connected with lights which flash on a board which may or may not be visible to the members themselves, depending on the experimenter's interest. The position of the lights can also be photographed at intervals during the meeting. Thelen has also hooked up the buttons with pens which provide a written record of member reactions.

Experience with this sort of equipment indicates that considerable training is needed in order for the subjects to remember to use the buttons or levers. Practice is also necessary to reduce interference with the group process. When there are complex response categories in terms of which the member must respond, these too must be learned and practiced. At the present time, there is no real evidence concerning the extent to which the use of these devices modifies the group process. Their maximal value would seem to accrue when used in connection with interviews of the participants and sound recordings. They are also valuable when the possible effect on the process itself is of minimal importance, as in training situations, for example.

Thelen and his associates have also done some work using physiological measures during group interaction. They have obtained G.S.R. and pulse rate records along with recordings. There are technical as well as theoretical complications in connection with collecting these kinds of data and, again, we have little evidence to indicate their usefulness.

DEVICES FOR PRESERVING A RECORD OF THE BEHAVIOR

We are concerned here with the fact that often it is desirable to obtain as complete a record of the behavior as possible. This may be necessary because observation is impossible. It may be desirable to check on the adequacy of on-the-spot observation by comparing it with categorizing or rating done later from the record. A permanent record also permits more analysis and study than is possible from a single exposure to the phenomenon. There is another occasion making a permanent record desirable. Sometimes it is impossible to make a judgment, or categorize a statement, until time has elapsed and the subsequent events have been noted. Thus, for example, a judgment of the relevance of a contribution might well have to wait until the final nature of the solution is known.

The most common device used to meet all or one of these objectives is that of *sound recording*. There are many recording devices now available, ranging in price and varying in their characteristics. By and large, it is possible to get adequate recordings, at least in terms of understandability, without special equipment such as soundproof rooms or other acoustical devices. Sound records have undoubted value in the study of social behavior, but the value is not without limits. The limits become more and more apparent the closer one moves toward dimensions of social interaction which depend on facial expressions, postural changes, and inflections of voice. The records often present identification problems when it is necessary to know who is speaking and to whom the remark is addressed. This difficulty can be overcome, in part at least, by having a carefully timed abbreviated stenographic record which contains identifying statements as well as the who-to-whom data.

Closely related to the use of sound records is the use of typescripts. These are verbatim typewritten records of verbal interaction taken from the sound record, typically with little interpretation on the part of the person making the typewritten record. These have had extensive use. Steinzor's category system, referred to above, was applied exclusively to typescripts. Experience has shown that typescripts are particularly useful in preliminary training of observers in the use of categories. There is unpublished evidence to indicate that when the category system involves very little inference, as in the case of the problem-solving categories above, there is no significant difference among the records obtained by observers who worked on the spot, by observers who worked directly from the recording, and by coders who worked from the typescripts.

One of the main advantages of the typescript is that it facilitates total recall of the situation and hence promotes re-entry into the data. The principal disadvantage is the cost and time involved in the typing process. Thelen (1950) reports that it takes from 7 to 10 hours to type out an hour of discussion.

Another possibility, although only rarely used up to now, is the use of *photography,* either motion picture or still. Thelen has used a 35mm camera which automatically photographed the actors at specified time intervals. This provides a periodic sampling of all overt behaviors such as locations of individuals, postures, facial expressions, activity of the group, changes in work group composition, and so on. Photographs have also been used by cultural anthropologists, notably Mead, to record characteristic features of social interaction in a society (Mead and Macgregor, 1951; Bateson and Mead, 1942).

Motion pictures have also been used to a limited extent. Lippitt (1940) photographed critical episodes in the life history of the groups exposed to different styles of leadership. Both still and motion pictures have primarily been used by researchers to document their observations or supplement data collected in other ways rather than as sources of data on which observations are made. There are real possibilities here, however, especially with motion pictures accompanied by a good sound track. Unfortunately, the expense involved in such a procedure makes it unlikely that these hypothesized values will get much of a chance to demonstrate their presence or absence.

Even a brief review of the devices that have been used in connection with observation of small groups indicates that their primary value lies in the extent to which they increase the skill, efficiency, and thoroughness of the human observer. In most cases they are aids for him, not substitutes.

There are many unanswered questions concerning the use of instruments. This is true for instruments which require direct use by the subjects themselves as well as those which are merely present and visible, as in the case of recording or photographing equipment. We do not yet know the extent to which they interfere or modify the nature of social interaction. More precisely, we do not know very clearly the conditions under which we should expect effects. We do know that there are circumstances, notably when the subjects become thoroughly accustomed to the equipment and its purpose is explained to them in a clear-cut, nonanxiety-producing way, under which the effects are apparently minimal.

Unfortunately, purely methodological research of this sort is not very exciting and is not likely to be done systematically. It is possible, however, to formulate methodological problems so that they deal with basic theoretical problems. This has been done partially in connection with the effect of observers on social interaction. When similar attempts are made in connection with devices, we may expect some answers to the questions which they pose.

PROBLEMS IN THE DEVELOPMENT OF AN OBSERVER SYSTEM

In this section we shall discuss some of the recurrent problems in the development of an observer system and its successful application to the data collection situation. This section is an extension of the earlier section concerned with dimensions of observer systems.

RELIABILITY

There is no need here to emphasize the importance of reliability as an attribute of a measuring instrument. Indeed, in the development of most of the systems described earlier a great deal of time was spent in efforts to arrive at reliable scores. Typically, reliability is defined in terms of degree of agreement between independent observers. The statistical indices which have been used to represent amount of agreement have varied considerably, but the most common has been the correlation coefficient. There are many circumstances where other indices are superior, but consideration of these questions falls outside the scope of this chapter. We propose here to discuss some general factors which influence the degree of reliability attained.

It seems useful to point out, at the outset, that there is only one score whose reliability need be assessed, and that is the score which is actually to be used in the analysis. Thus, for example, when group members need only be ordered in terms of the extent to which they showed a particular behavior during a meeting in order for a hypothesis to be tested, it is the reliability of the total scores for each individual which must be determined. It often happens that the reliability of this score is quite high, even though the degree of agreement between observers in their categorization of single units of behavior is much lower. This typically happens when differences in coding of single units are randomly distributed among the categories.

The general evidence on reliability of observer scores indicates that the less inference required of observers, the higher the degree of agreement. With category or rating systems which require a good deal of inference, it is clear that the most reliable are those in which the dimensions are clearly defined and the cues to be used by observers are specified. The observational load is a factor affecting reliability. Other things being equal, the more the observer has to do, the lower the reliability. Some preliminary results of work done by students of the authors indicates that there is less of a decrement in reliability when additional observations require the same level of inference than when the same number of categories are added which require more inference than the original set.

The reliability of some systems is low because of unclear definition of the unit. Even when the unit to be rated or classified may be fairly clearly stated, the observers may nevertheless disagree as to unit boundaries in actual observing situations. It is obvious that degree of agreement as to coding can only be appropriately assessed when there is agreement first as to what is to be coded.

Up to the present time, particularly with category or rating systems which require some inference, the degree of reliability attained is very much a function of the amount of training which observers have had. The thoroughly trained observer has become maximally familiar with the definitions of the categories or dimensions. He has gotten practice in designating the unit of observation. He has gained skill in adopting the frame of reference which the system demands, and has practiced his art in the situation in which the behavior takes place. His own private definitions have been eliminated or have come to be shared by others using the system. In short, virtually all the sources of unreliability have been dealt with in the training process. Lippitt and his associates have found that this sort of training is maximally effective when (1) the observers assist in the development of the observer system, and (2) the training is concentrated on correct classification of marginal or difficult units rather than upon items which clearly fall into a given category.

Ideally, every observational system ought to be developed to the point that it can be communicated to other social scientists and applied reliably by them. Only in this way can we really say that an instrument has been developed. This does not mean that the instructions, definitions, and procedures must be simple; it does mean that they should be readily shared, and with a minimum of ambiguity. Unfortunately, this state of affairs is not yet with us. The typical experience is for one experimenter to try the observational procedures of another and then give up the attempt because of a failure to attain sufficient reliability. This has not always been due to insufficient *clarity* of presentation. There is often a failure to communicate some of the more subtle aspects of the procedure which are a part of the technique of the originator. It is also quite likely that some of the failures to reproduce the success of the developer of the system have been due to insufficient training time. There are certain factors which affect the reliability of sources which no amount of specification will enable the observer to handle. The observer must have the chance to develop sensitivity to the dimensions of group process under scrutiny. He must learn to cope with the speed of interaction which so frequently overwhelms the inexperienced. To deal successfully with the speed of the interaction process, the category system itself must no longer be strange.

Because of the difficulties involved in obtaining satisfactory reliability and the responsiveness of reliability scores to training, it is virtually mandatory that reliability checks be run prior to securing the research data. Reliability figures obtained after considerable practice may have no bearing at all on the reliability of the original scores.

VALIDITY

It has been pointed out by Cronbach (1946), Guttman (1950), and others that the term *validity* has a variety of meanings. In one sense, the validity question means: Does the observer score measure what it purports to measure? In another sense, the question of validity asks: Does the observer score predict anything? If we accept the first meaning of the term, the question of validity of observer scores has been relatively ignored by those people most active in using observers.

One of the reasons for the neglect has been that often the categories require very little in the way of inference. They have what has been called "internal validity" (Guttman, 1950). No one seriously doubts that the score in the category "asks question" represents what it is supposed to represent, least of all the people using the system. The definition of the category is universally shared and there is common acceptance that the average person knows a question when he hears one.

The problem is much more complex in the case of category systems which involve inferences as to intent or motive or feeling of the actor or recipient. It is a quite legitimate question to ask whether the entries in the "feels hostile" category really represent the degree of hostile feelings present. It is rare, however, that a really satisfactory independent measure of the same variable is available with which the observer scores can be compared. It is this difficulty which has prevented much work on the validity of observer scores.

There is at least one solution to the problem of determining the validity of an observer score in the absence of an adequate external, more acceptable, criterion score. This solution lies in determining the extent to which the scores in question behave as they are supposed to behave as determined by a well-formulated and documented theoretical position. To illustrate, let us take a study reported by Thibaut and Coules (1952). These experimenters were concerned with the effects of variation in opportunity to communicate hostile feeling on the amount of residual hostility. To determine the relationship, they coded the content of written documents into categories having to do with the nature of the feeling and the referent of the feeling. Let us suppose for purposes of our illustration that the proposition that opportunity to communicate directly to the instigator of frustration reduces the incidence of hostile responses had already been supported by other studies, using other assessment techniques. Then this same design could be used, because of the support independently provided for the theory, to test the validity of the content coding. This same thing could, of course, be done with observer scores. Unfortunately, it is difficult to find such well-articulated and documented theory which can be used to validate an observational procedure. Nevertheless, this procedure is often more likely to prove useful than a search for an adequate independent measure of the same variables.

If we turn to the second meaning of the term *validity*, the picture is much brighter. There is no need to enumerate all of the instances in which observer scores have been found to correlate with other measures. It has been possible to achieve statistically significant relationships between participant satisfaction with meetings and observer ratings of the cohesiveness of the group, and the amount of self-oriented need behavior, for example (Guetzkow, 1950). Observer scores distinguish the behavior of group members subjected to different kinds of leadership, and so on. Indeed, it is the promise of these kinds of results which is responsible for the energy which is now being expended in the development of observer systems. In spite of this general picture of moderate usefulness of observer scores, it cannot be said that there is evidence that all of the systems which have been used have such predictive efficiency. It seems fair to say that those which have such support tend to be of the higher-inference sort. The experience of Thomas (1933a) and her associates indicates that it is possible to get extremely high reliability about unimportant matters. This would suggest that mere demonstration that the system can be used reliably is not sufficient warrant for publishing it abroad. Evidence should be provided quickly that the procedure is useful in prediction of some sort.

SAMPLING

The question frequently arises as to whether the observer should be asked to record all of the behavior in a given period or whether some sampling should be specified. There is a wide variety of ways in which selection can be made from the total behavior period. The observer can systematically rotate his attention among all the members of the group, making each one the subject of observation for a specified period. Or he can select certain members and ignore others. Another possibility is to observe all members, but at specified times, rather than continuously. There are many others.

All sampling problems, to be handled successfully, require a knowledge of the universe from which the sample is to be drawn. They also require, under ideal conditions, that the theory guiding the research be well formulated. At the very least, it is necessary to be clear about the assumptions one is making when one samples in any of the possible ways, and the effects on the data if these assumptions are not met. Unfortunately, we know so little about social interaction that sampling procedures are virtually always risky. In fact, one of the principal objectives of developing procedures of social observation is to describe social interaction. It would seem desirable to make these procedures as all-encompassing as possible at the outset and leave questions of sampling until we have more empirical data on the basis of which they can be settled definitively.

This is not to say that there are not conditions under which the investigation can be carried out successfully with sampling of people or of time or both. Anderson (1939) reports, for example, that there is a high degree of consistency in teacher behavior toward children from one time period to another, indicating that sampling, using his system, is probably quite justified. In these cases, however, it is the theory and the design which dictate whether sampling is sound in the circumstances, and not propositions about sampling in general.

PRE- AND POSTCATEGORIZATION

Sometimes there is the possibility, in theory at least, of securing a complete record of the total behavior and doing the categorizing job on this recorded material. In this sense, postcategorization refers to coding (or rating) from a record rather than at the time the behavior takes place. Whether this is an adequate procedure depends on the extent to which correct use of the categories necessitates cues which are not available because of the type of record used. Thus, for example, typescripts (typewritten records of interaction, primarily verbal) typically do not provide information concerning facial expressions, voice inflections, postural changes, which often play an important role in determining judgments concerning emotional states. When the category system requires little use of such cues, the results obtained with on-the-spot coding and afterward from the typescript are very similar.

There is another sense in which the term *postcategorization* is used. This second meaning refers to the practice of *developing the category system* in the process of studying some sort of total record. This seems to us to be almost entirely a historical question in the case of instruments which make some claim to being standardized. In these cases, there is an extensive history of total observation and total recording and the progressive refinement of the category system. The initial step is typically that of deriving the categories in part at least from the behavior itself. At various times in this sequence, categories are added or dropped on the basis of what types of behavior appear and can be differentiated. There are, therefore, times in the life history of most observation schedules when they are essentially postcategorization systems. Indeed, any system which purports to be applicable to a wide variety of social situations must have that sort of history if its claim to such generality of use is to have any validity at all.

THE EFFECT OF OBSERVERS ON THE BEHAVIOR ITSELF

This is a problem which perturbs the nonuser of observation techniques much more than it does those who have worked with them extensively. There are undoubtedly many reasons for this differential interest in the issue, but certainly one of them is the common feeling among experienced users of observers that the observers have very little effect, if any. This belief is shared by experimenters who have worked in a wide variety of situations and with many different kinds of subjects.

While this general belief is probably sufficient to encourage most people to go ahead constructing observation schedules and using observers, it is also true that our present state of knowledge on this issue is very inadequate. Whether or not the presence of the observer has an effect on the nature of the interaction depends undoubtedly on such factors as the nature of the group, the nature of the activity of the group, and the way the observer task has been structured to the group members. Unfortunately, there has been little systematic study of the relationship between these factors and the effect of observers. To supplement the following brief discussion the reader should turn to Chapter 21, where this problem is also treated.

Deutsch (1949) reports that over a period of three weeks members of small groups became less and less aware of the presence of observers, until by the end of the period they were virtually oblivious of them. In certain of these groups the activity involved cooperation among members, while in others the members were competing with each other. The competitive groups were more aware of the observers than were the cooperative groups. This suggests, as indicated above, that the nature of the activity is a factor affecting awareness of observers if not the social behavior itself.

In contrast to Deutsch, Polansky *et al.* (1949) cite some experience which suggests that there are circumstances in which subjects being observed become more aware and less accepting of observers over time. Their subjects were boys in a summer camp. At the outset, the boys accepted and tended to ignore the observers, who, they had been told, were trying to see how to improve the camp. By the third week, the observers were objects of attack by many of the boys. Polansky suggests that the ambiguity of the role of the observer made the boys anxious. When the observers became more friendly to the boys, the attacks were less frequent. This study indicates the importance of having the role of the observer clearly defined but defined in such a way as to have the minimum impact on the behavior itself. Thus, for example, it seemed desirable to indicate to conference groups in business organizations that the observers had no relationship to management, that no one was personally identified, and that the sole purpose of the observer team was to record what took place in the interest of finding out how business conferences were conducted. While this structuring was only part of the total procedure, it is interesting to note that almost all the members in 72 groups indicated that they felt the presence of the observers had had no effect on what had taken place. Since these were groups which had a history of previous meetings, they were able to compare the meeting at which the observers were present with the earlier ones.

The Ideal Conditions: Analogy to Psychophysics

Thurstone has explicitly applied traditional psychophysical methods such as paired comparisons and equal-appearing intervals to the measurement of attitudes. As we mentioned at the outset, the relationship between human reactions and independently assessed stimulus properties has been most thoroughly investigated in psychophysics. It seems profitable for us to search this area for implications for the use of observers in social psychological research.

By and large, experimental results in psychophysics are very encouraging, since they show that the human being can make judgments about external events which bear an amazing correspondence to the relationships between these events as determined by independent techniques. The psychophysicists have also been able to isolate some of the conditions which affect this degree of correspondence. These findings can be grouped under four major headings:

1. When other conditions are optimal, the human being is more accurate when he makes ordinal judgments than when he makes judgments on an absolute subjective scale. Almost all psychophysical techniques require the judge to say whether one object is more or less of something than another. In general, human beings have shown themselves to be very sensitive to ordinal differences between objects which can be shown to correspond to ordinal differences determined by independent techniques.

2. Ordinal judgments are more accurate when there are only two objects to be judged at a time. The greater the number of objects presented simultaneously for ordering, the more inaccurate the judgments.

3. Ordinal judgments on two objects are most accurate when the objects of judgments are brought close together in space and time. With lifted weights, the longer the interval between the lifting of the standard and the comparison weight, the more inaccurate the judgment. In the judgment of the relative length of two lines, the smaller the space between them, the more accurate the ordinal judgment — up to a point at which they can no longer be distinguished as two lines.

4. Ordinal judgments of two objects close together in space and time are most accurate when the objects are identical on all attributes except those being judged. Judgments of relative length are most accurate when the two lines are the same width.

Psychophysicists are in the fortunate position of having an independent criterion for the accuracy of psychophysical judgments. For most of the attributes of interest to the social psychologist there is no such independent criterion. Nevertheless, it seems reasonable to assume that these same principles, if applied to social psychological judgments, should produce maximal correspondence to the objects of judgment. In any event, it is significant that these conditions are the very ones in which the human instrument experiences least frustration and uncertainty in performing his task. This should be an important consideration, since we do not want to damage this instrument any more than we have to.

Let us suppose that we wish to apply these principles derived from psychophysics as directly as possible to the gathering of data for the study of the decision-making conference. While the following discussion is written in terms of assigning scores to groups as units, it should be noted that it applies as well to conditions where individuals or periods of time are being judged. Let us also suppose that we have unlimited time and money at our command. With slavish adherence to these principles, what would the methodological procedures be?

> *Ordinal judgments are more accurate than absolute judgments.* This would mean that we allow our observers only to rate "more" or "less," never allowing them to assign numbers to their ratings.

The psychophysicists have found that such ordinal judgments are most accurate under optimal conditions. However, as one departs from these conditions, ordinal judgments become more and more inaccurate. Inter- and intraobserver variability rises. In the size constancy experiment, for example, if the two objects to be compared are so placed that one is in front of the observer and one behind, the observer will have to turn around for each judgment. This condition will produce a greater percentage of judgments which do not correspond to the "physical" relationship between the two objects.

Now the conditions of social psychological measurement in the field are often far from optimal in this respect. There may be a prolonged period between observation sessions. When such conditions become bad enough, it may be that there comes a point at which ordinal judgments become more variable than absolute judgments, that is, conventional numerical ratings. It is conceivable that absolute judgments have a large but finite variability which is less than the variability of ordinal judgments when conditions of judgments depart far enough from the optimum. This is a matter for empirical demonstration, and if it turns out that there is such a point of diminishing returns, it will be a matter of decision in planning a study as to which method will be most accurate under a given set of conditions.

When we consider what the task of the observer becomes when we depart from optimal psychophysical conditions, the rationale for this point of diminishing returns becomes obvious. The conditions are optimal when the observer can have both objects in his field at once so that he can quickly shift from one to the other in making his comparison. When the objects are moved far apart in time or space, it means that the observer must "remember" the first object while he is turning to the comparison object. We are, in effect, asking the observer to compare a *memory trace* of the first object with the immediate experience of the second. With a longer interval, it is likely that this trace undergoes some changes; and there is no reason to assume that these changes are in a constant direction. The traces are probably affected to a large extent by what else happens in the interval, and the things which happen to one observer may be very different from the things which happen to another. Consequently, we would expect there to be more disagreement between observers as the interval between standard and comparison object is increased. Exactly the same argument holds for repeated comparisons of one observer.

Now what is the situation when we ask an observer to make an absolute judgment, a rating? Essentially, the task is the same as the one we have just described. We are asking him to compare an object in his immediate field with a trace which has been built up over a long period of time, a trace corresponding to "inches" or "brightness." If the memory trace of "inches" has been built up from a large number of experiences, one would expect that it will be affected relatively little by the additional information given during the judgment. If the trace is a new one or one which has been established on the basis of relatively few experiences, we should expect it to be affected to a much greater degree by new experiences in the process of judging. The subjective scale for such an attribute would be much more variable, and consequently we should expect that we could depart from the optimal psychophysical conditions a great deal more and still have ordinal judgments more accurate than absolute judgments. Again this is a matter for empirical determination. It would be easy to determine how variable a subjective scale is. In the light of this, one could have some criterion in terms of which to select the more appropriate method of judgment under a given set of conditions.

> *Limit the objects of judgment to two.* Observers should make all their judgments with reference to one other group, person, or time period, and this other group, person, or time period should be the same for all observers.

As a matter of fact, this is what is done ordinarily, whether the observers are asked to give relative or absolute judgments. If the factors described above point to the advisability of using ordinal judgments, care should be exercised in the study design to make sure that the observers make their comparisons with reference to one other meeting, and that this reference conference is the same for all observers at any one time of judgment. If the situation makes absolute ratings seem advisable, then care must be taken to make the subjective scales of the observers as congruent as possible and to assure that the scales do not change with the passage of time. This is ordinarily done by defining a reference group, real or hypothetical, which is used as a base line for all ratings.

> *Proximity of objects of judgment in space and time.* The pairs of groups, persons, or time periods to be compared should be as close together as possible.

This is the condition which will probably be most difficult to fulfill. The judges cannot ordinarily schedule the groups to suit their needs. Obviously, the judges can't observe two groups simultaneously, but the scheduling of groups to be observed should be as close as possible. As has been pointed out above, this is one of the conditions which will determine which type of judgment will be most appropriate for a particular study. If long periods of time must intervene between observation sessions, then absolute judgments (conventional numerical ratings) may contain less error than ordinal comparison judgments. If this is so, care must be taken to keep the reference norms for such ratings fresh in the judges' minds. The object is to make the conditions approximate as closely as possible those of simultaneous comparison.

> *Similarity of objects of judgment on all attributes not being judged.*

If accuracy of measurement is the main consideration, this rule has obvious implications for the selection of a sample of groups or persons to be observed. If the period of time elapsing between group sessions were not a limiting factor, then the way to achieve this condition would be by observing one conference on successive meetings. This procedure would automatically keep many factors constant which would be almost impossible to control in any other way. With this technique we would have our observers make their comparisons against the previous meeting, since this would be the one which was most "fresh" in their minds. Successive comparisons would be made on a sufficiently long series of meetings so that tests of significance could be made on the data so obtained.

If it were found possible to do a study under these ideal psychophysical measuring conditions so that one had a series of ordinal judgments on successive meetings of one conference group, what sort of analysis could be done with the data so obtained? If more than one observer made judgments at each meeting, it would be possible to arrive at a conventional index of reliability on each attribute. With such a series of "more" or "less" judgments, we would be able to detect changes in an attribute which were monotonic, that is, changes which showed increase or decrease throughout the series of meetings. If the changes in an attribute were nonmonotonic, our data would show it, although we would not be able to specify the shape of this nonmonotonic function. The data would provide a basis for stating relationships between attributes. These statements would be functionally comparable to conventional correlation coefficients, but would not require either the more inaccurate methods of data collection, or the strenuous statistical assumptions needed for the computation of a product moment coefficient. We could also compute a sampling error for the obtained relationships.

These data, it is evident, offer about as many possibilities for analysis as do the conventional numerical ratings, although no index can be computed from them which requires the assumption of a distribution function. This is not a serious limitation, since there is a rapidly growing set of statistical devices which involve no assumptions about distribution functions. (See Chapter 8.)

Undoubtedly this movement of a research design in the direction which makes observer data more satisfactory is bought at the cost of a reduced generality. The results of a study of a hundred different groups can be generalized more safely than the results obtained from a study of a single group one hundred times. This is unquestionably true. There are, however, other ways in which this optimum condition of similarity on the other attributes can be approached without a material loss in generality. In many cases, these dissimilarities can be eliminated in the final design by selection of subjects; in still others, these influences can be minimized by making their effects an explicit part of the training procedure.

The implications of the results of psychophysical experimentation have been developed in the above discussion in a setting of observers who were called upon to make judgments of group properties. Each of them can with equal merit be applied to observation of individuals. Indeed, as in the case of optimal conditions 2 and 3, the task of implementing them is easier when the subject of observation and analysis is the individual.

Finally, it should be noted that these conditions of optimal use do not eliminate the need for careful definition of the dimensions of observation and the equally thoughtful examination of the other factors affecting successful use of the observer which have already been mentioned. The discussion of the implications of psychophysics calls attention to possible criteria in terms of which to evaluate possible procedures for the use of observers.

CONCLUSION

The observer is now an established member of the social science research team. His services have been called upon in a wide variety of situations. He has done everything from counting the number of times a person says "We" to making judgments about motivation. By and large, the observer has earned the attention and respect he is currently receiving. The data he provides are increasing our understanding of complex social psychological phenomena. In the instances where observation has failed to provide useful data, the difficulties are usually found not in the individual observer but rather in the design of the study, the failure to define clearly the area of observation, and inadequate training.

There is a rapidly growing library of observational systems applicable to many areas of social behavior. They are becoming increasingly sophisticated, methodologically. One of the most serious barriers to scientific advance has been the fact that these systems tend to be used only by their authors. Other researchers working in the same areas are prone to develop their own procedures. The result is that the problem of comparability of findings is complicated. It seems reasonable that some of the systems now available, with respectable histories, ought to be used extensively in much the same way that established tests and standard learning tasks are used in other areas of social science.

Finally, there are many important methodological questions still to be answered. This is primarily because almost no one who has worked in the area has been primarily a measurement methodologist. Most of the methodological sophistication has resulted from being forced to confront technical problems en route to finding answers to theoretical questions. In spite of the lack of strictly methodological experimentation, there is an increasing awareness of the critical questions which must be asked in the development of an observer system. There is, in the history of systems already developed, much to guide the scientist who feels obliged to develop his own.

REFERENCES

ALLPORT, G. W. *Personality*. New York: Holt, 1937.

ANDERSON, H. H. Domination and social integration in the behavior of kindergarten children and teachers. *Genetic Psychol. Monogr.*, 1939, *21*, No. 3, 285-385.

BALES, R. F. *Interaction Process Analysis*. Cambridge, Mass.: Addison-Wesley, 1951.

BALES, R. F., & GERBRANDS, H. The interaction recorder – an apparatus and check list for sequential content analysis of social interaction. *Hum. Relat.*, 1948, *1*, 456-464.

BARKER, R. G., & WRIGHT, H. F. *One Boy's Day*. New York: Harper, 1951.

BATESON, G., & MEAD, MARGARET *Balinese Character, a Photographic Analysis*. New York: The New York Academy of Science, 1942.

BERGMAN, D. V. Counseling method and client responses. *J. consult. Psychol.*, 1951, *15*, 216-224.

BÜHLER, CHARLOTTE *The first year of life*. New York: John Day, 1930.

BÜHLER, CHARLOTTE Social behavior of the child. In C. Murchison (Ed.), *A Handbook of Child Psychology*. Worcester, Mass.: Clark University Press, 1931. Pp. 392-431.

CARTER, L., HAYTHORN, W., MEIROWITZ, BEATRICE, & LANZETTA, J. The relation of categorizations and ratings in the observation of group behavior. *Hum. Relat.*, 1951a, *4*, 239-254.

CARTER, L., HAYTHORN, W., MEIROWITZ, BEATRICE, & LANZETTA, J. A note on a new technique of interaction recording. *J. abnorm. soc. Psychol.*, 1951b, *46*, 258-260.

CHAPPLE, E. D. Measuring human relations: an introduction to the study of the interaction of individuals. *Genet. Psychol. Monogr.*, 1940, *22*, 1-147.

CHAPPLE, E. D. The interaction chronograph: its evolution and present application. *Personnel*, 1949, *25*, 295-307.

CHAPPLE, E. D., & LINDEMANN, E. Clinical implications of measurements of interaction rates in psychiatric interviews. *Applied Anthrop.*, 1942, *1*, 1-11.

COFFEY, H. S., FREEDMAN, M. B., LEARY, T. F., & OSSORIO, A. G. Community service and social research – Group psychotherapy in a church program. *J. soc. Issues*, 1950, *6*, 1-65.

CRONBACH, L. J. Response sets and test validity. *Educ. psychol. Measmt.*, 1946, *6*, 475-494.

DEUTSCH, M. An experimental study of the effects of cooperation and competition upon group processes. *Hum. Relat.*, 1949, *2*, 199-231.

DITTMAN, A. T. The interpersonal process in psychotherapy: Development of a research method. *J. abnorm. soc. Psychol.*, 1952, *47*, 236-244.

FOURIEZOS, N., HUTT, M., & GUETZKOW, H. Measurement of self-oriented needs in discussion groups. *J. abnorm. soc. Psychol.*, 1950, *45*, 682-690.

FREEDMAN, M. B., LEARY, T. F., OSSORIO, A. G., & COFFEY, H. S. The interpersonal dimension of personality. *J. Pers.*, 1951, *20*, 143-161.

GHISELLI, E. E., & BROWN, C. W. *Personnel and Industrial Psychology.* New York: McGraw-Hill, 1948.

GUETZKOW, H. S. *Groups, Leadership and Men.* Pittsburgh: Carnegie Press, 1951.

GUTTMAN, L. The problem of attitude and opinion measurement. In S. A. Stouffer, *et al.*, *Measurement and Prediction.* Princeton: N. J.: Princeton Univ. Press, 1950. Pp. 46-59.

HEYNS, R. W. Functional analysis of group problem solving behavior. Ann Arbor, Michigan: Conference Research, Department of Psychology, University of Michigan, 1948 (mimeographed).

JACK, LOIS M. An experimental study of ascendant behavior in preschool children. *Univ. Iowa Studies in Child Welfare*, 1934, *9*, No. 3.

KEET, C. D. Two verbal techniques in a miniature counseling situation. *Psychol. Monogr.*, 1948, *62*, No. 294.

LIPPITT, R. An experimental study of the effect of democratic and authoritarian group atmospheres. Studies in topological and vector psychology. I. *Univ. Iowa Studies in Child Welfare*, 1940, *16*, No. 3.

LIPPITT, R., & ZANDER, A. Observation and interview methods for the Leadership Training Study. New York: 2 Park Avenue, Boy Scouts of America, 1943 (mimeographed).

MEAD, MARGARET, & MACGREGOR, FRANCES C. *Growth and Culture.* New York: Putnam, 1951.

OLSON, W. C., & WILKINSON, N. M. The measurement of child behavior in terms of its social stimulus value. *J. exp. Educ.*, 1932, *1*, 93-95.

OSBURN, H. An investigation of the ambiguity dimension of counselor behavior. Unpublished doctor's dissertation, University of Michigan, 1951.

POLANSKY, N., LIPPITT, R., & REDL, F. Problems of interpersonal relations in research on groups. *Hum. Relat.*, 1949, *2*, 281-292.

POLANSKY, N., LIPPITT, R., & REDL, F. An investigation of behavioral contagion in groups. *Hum. Relat.*, 1950, *3*, 319-348.

ROGERS, C. R. *Counseling and Psychotherapy.* Boston: Houghton Mifflin, 1942.

SNYDER, W. U. An investigation of the nature of non-directive therapy. *J. gen. Psychol.*, 1945, *33*, 193-223.

STEINZOR, B. The development and evaluation of a measure of social interaction. *Hum. Relat.*, 1949, *2*, 103-122.

THELEN, H. Techniques for collecting data on interaction. *J. soc. Issues*, 1950, *6*, 77-93.

THELEN, H., & WITHALL, J. Three frames of reference: a description of climate. *Hum. Relat.*, 1949, *2*, 159-176.

THIBAUT, J. W., & COULES, J. The role of communication in the reduction of interpersonal hostility. *J. abnorm. soc. Psychol.*, 1952, *47*, 770-777.

THOMAS, DOROTHY S. An attempt to develop precise measurement in the social behavior field. *Sociologus*, 1933, *9*, 1-21.

THOMAS, DOROTHY S., LOOMIS, A. M., & ARRINGTON, R. E. *Observational Studies of Social Behavior.* Vol. I. *Social Behavior Patterns.* New Haven, Conn.: Yale University (Institute of Human Relations), 1933.

WRIGHT, H. F., & BARKER, R. G. *Methods in Psychological Ecology: A Progress Report.* Topeka, Kansas: Ray's Printing Service, 1950.

CHAPTER 11

Sociometric Measurement

GARDNER LINDZEY AND EDGAR F. BORGATTA*
Harvard University

The web of interpersonal relations, the attractions, repulsions, and indifferences that characterize individuals in daily interaction, the informal organization of groups, the social status of individuals — all these are of interest to the social psychologist, and aspects of all may be measured effectively through the use of sociometric techniques. These instruments are designed specifically to provide a sensitive and objective picture of the interpersonal relations existing within a group. Consequently, they are of singular importance to the empirically oriented social psychologist, whether he is interested in theoretically derived or applied research.

Use of the term "sociometry" has been surrounded by considerable disagreement and debate. Its similarity to "biometrics" and "psychometrics" suggests that it might refer to all measuring devices that deal with social behavior. Consistently with this, Moreno (1943a) has urged that sociometry should not be limited solely to instruments he has developed. Moreover, Bain (1943) indicated that the term properly should refer to all attempts to measure social behavior, and Chapin (1943, 1940), Lundberg (1943), and Sanderson (1943) express general agreement with this opinion. On the other hand, the introduction and persistent use of the term by Moreno has led to common association in the minds of most social scientists between "sociometry" and *only* those techniques introduced by Moreno. In modern texts of social psychology the term is used without exception in the more restricted sense. Thus, authorities argue one thing, common usage another. We shall bow to common convention in our discussion and consider the term *sociometry* appropriate to designate a limited range of measuring devices, all of which stem primarily from the work of Moreno.

Moreno (1934) refers to the measures with which we are concerned in this chapter as sociometric tests. He reserves the more general term, sociometric techniques, to refer to a cluster of devices, including the spontaneity test, the various types of role tests, and the acquaintance test, in addition to the sociometric test. Although we use the terms test, technique, and measure interchangeably, our interest will be limited to those measures Moreno designates as sociometric tests. Further, the scope of this chapter prohibits consideration of the theoretical formulations which have been commonly associated with sociometry and Moreno.

The focus of our interest is upon a small number of closely related techniques that may be used to elicit responses from members of a given group concerning the positive, neutral, and negative relations existing within the group. In the following sections we shall consider the conditions under which such responses are secured, the primary ways in which the resultant data may be analyzed, some common properties found in the analyzed data, and the areas of investigation in which such data may be expected to be of significance.

What qualities do sociometric measures possess that make them of special interest or significance to the social psychologist? While the detailed answer to this question is contained in the body of this chapter, there are a number of general points that may be suggested at this juncture. *First,* it is evident that the variables these measures represent are about as purely "social" as is possible. The fact that these variables describe the quality of interpersonal relations within groups makes clear that they deal with the basic data with which the social psychologist has been most interested, both conceptually and empirically. *Second,* the interdisciplinary popularity and appropriateness of these devices makes them of particular relevance to a subdiscipline that has developed

---

*Much of the work on this chapter was carried out as part of a program of research on interpersonal perception at Harvard University, financed by the Office of Naval Research (Contract No. N5-ori-07646). We are deeply indebted to many colleagues whose comments have helped in shaping this chapter. Especially helpful were the suggestions of Joan H. Criswell, Mary L. Northway, and Charles Proctor.

largely in the void existing between two parent disciplines. A striking feature of social psychology during the past decade is the increase in interdisciplinary interest and multidiscipline research. Such an orientation may be expected to have a fruitful outcome to the extent that individuals in the different disciplines are able to agree upon the mutual relevance and significance of data. In this connection, it is interesting that from the very beginning sociometric techniques have been widely employed by both sociologists and social psychologists, and to a lesser extent by social anthropologists and psychiatrists. The development of these measures is associated with no single discipline but, rather, appears to have occurred while many disciplines have had access to their use. Consequently, sociometric devices seem ideally equipped to function in a field where there is much interest in cross-disciplinary integration. *Third,* and this is related to the interdisciplinary popularity of these techniques, is their capacity to represent individuals in interaction within a miniature social system. Current theoretical formulations, in addition to the demands imposed by many empirical problems, make it necessary for the investigator to study the individual and his social environment simultaneously. The sociometric measures provide a means of representing an important part of the individual's social environment as it is perceived *by the subject.* Recent emphasis on the importance of the "subjective frame of reference" or the "psychological environment" is thus satisfied by these techniques. Further, they provide an externalized view of the same events through the responses of other members of the group. *Fourth,* ease and speed of administration and a related lack of expense are important qualities in an era where much investigation has become so expensive that only the wealthy or well-endowed may hope to compete successfully. These are devices that the independent researcher can use effectively without large-scale resources. *Fifth,* increasing emphasis upon action research, investigation where the findings have direct implications for the concrete events under study, gives further merit to sociometric techniques, as they lend themselves to the introduction of social change in an efficient and compelling fashion without seriously adding to the labor of the investigator. *Sixth,* sociometric measures can be used in such a way as to provide a much higher degree of interest and motivation on the part of participant subjects than is typical of most psychological measuring instruments. The possibility that his environment may be manipulated so as to comply with his wishes encourages a high degree of interest and cooperativeness on the part of the respondent. *Seventh,* as we shall attempt to document later, sociometric techniques provide relatively acceptable indices for a large number of empirical concepts that the social psychologist conventionally employs. Such concepts as prejudice, morale, social status, rejection, etc., can be measured effectively through the use of these techniques.

What other instruments in the psychologists' measurement battery show similarities to sociometric measures? Although there are some correspondences between these measures and attitude scales as well as questionnaires or inventories, the most important congruence seems to exist in relation to rating scales (P. N. Pepinsky, 1949; Jennings, 1950). It is easy to conceive of the sociometric test as a variety of rating scale — the members of the group are asked to rate or order the other members of the group in terms of their attractiveness or desirability for sharing certain activities.

In spite of an underlying similarity, there are also important differences between these two approaches. Sociometric measures are more limited in the variables they can be used to assess and more restricted in the settings where they can be employed. On the other hand, there is no need to train raters to engage in sociometric ratings. The difficult and time-consuming task of attempting to produce common frames of reference and homogeneous criteria in terms of which ratings shall be assigned is avoided. The rater is asked to apply exactly those particular, unique, and sometimes irrational criteria he has spent a lifetime developing. Everyone is an experienced or expert rater when it comes to sociometric judgments. Each of us has a vast body of experience in deciding with whom we wish to interact and whom we wish to avoid. Liking and disliking, accepting and rejecting are part of the process of daily living. Given the past experience of these raters and the meaningfulness of the activity they are engaged in, it is small wonder sociometric devices have proven a powerful measurement tool in the hands of the social scientist. One might say that the individual who uses these techniques is taking advantage of the largest pool of sensitive and experienced raters that is anywhere available.

ORIGINS OF SOCIOMETRY

The origins and development of sociometry are intimately linked with one man — J. L. Moreno. There are few instances where a single individual has exerted so pervasive and crucial an influence on the evolution of a social science area. Although it is clear that similar devices were used by other investigators prior to Moreno and that Moreno himself had worked with these devices before they became widely known, the important first step in the development of sociometry was the publication of *Who Shall Survive?* (1934). This volume reports a series of group studies carried out in public school and institutional settings using a variety of data-gathering techniques, including sociometric measures. The book is not a simple report of research, for it contains a mixture of speculative theory and applications of methods and their results to problems of broad social significance. This fusion of procedural skill, speculative theorizing, and heavy concern with evaluative issues has remained typical of much sociometric writing up to the present. Following this initial presentation of the method, the next important event was the publication of the short-lived *Sociometric Review* (1936). This was superseded by *Sociometry: A Journal of Interpersonal Relations,* the first volume of which appeared in 1937-1938. This periodical has played a crucial role in subsequent developments. Although the journal has been generous in publishing a large number of studies that bear no intimate relationship to Moreno's work, the over-all impression is certainly one of an official or home organ. This has led to more of an air of sectarianism than might otherwise have appeared. In spite of the fact that the first two editors of the journal were respectively a distinguished psychologist, Gardner Murphy, and a distinguished sociologist, George Lundberg, the expansive, persuasive influence of Moreno was so great that most psychologists and sociologists tended to look upon individuals concerned with sociometry as having some of the cultish qualities that were characteristic of Rorschachers and psychoanalysts. This set of conditions may have served to slow down the interpenetration of conception and method in the various areas, but at the same time it performed the important function of preventing sociometric techniques from becoming seen as tools that were the primary property of either sociology or psychology.

Whatever tendencies to set sociometrists off as a group by themselves may still remain, it is clear by now that sociometric measures are widely used by social psychologists and sociologists, and that some knowledge of these techniques would be expected of any trained individual in either of these areas.

WHAT IS A SOCIOMETRIC MEASURE?

In simplest terms, a sociometric measure is a means of assessing the attractions, or attractions and repulsions, within a given group. It usually involves each member of the group privately specifying a number of other persons in the group with whom he would like to engage in some particular activity and, further, a number of persons with whom he would *not* like to participate in the activity.

The most important requirements of the sociometric test generally advocated by Moreno are:

(1) The *limits of the group should be indicated* to the subjects. The sociometric test places no restrictions on the persons *within* the group who may be chosen or rejected, but subjects should clearly understand the nature of this group.

(2) The subjects should be permitted an *unlimited number of choices* or rejections. They should be encouraged to choose as many or as few of the group members as they wish.

(3) The subjects should be asked to indicate the individuals they *choose* or *reject in terms of specific criteria.* Each sociometric choice or rejection should be made with a particular activity in mind, and the activity should be meaningful to the subjects.

(4) Results of the sociometric questions should be used to *restructure the group.* Moreover, the subjects should be told that their choices and rejections will play a decisive role in determining with whom they will associate in this activity.

(5) The subjects should be permitted to make their choices and rejections *privately,* without other members of the group being able to identify the responses.

(6) The *questions* used *should be gauged to the level of understanding of the members of the group.* With older children and adults, the questionnaire form of presentation is usually appropriate. With younger children who are not yet facile with writing, individual reports to the administrator of the sociometric test may be most appropriate. With the very

young and others who do not have adequate control of language, responses to a sociometric question may be meaningless, although other instruments may function effectively.

Except in cases where special empirical interests prohibit it, members of the group should have been in association for a sufficient time to permit the formation of affective ties and repulsions within the group. If stability of choice and rejection is important, it is evident that this cannot be expected until the members of the group have shared experiences over a period of time. Again, if it is desired that the members of the group should have had an opportunity to interact with all other members, the group should be small enough to permit this. Lastly, the group should be so constituted that it will be possible to restructure the group in the light of the sociometric findings.

VARIATIONS IN PROCEDURE

The requirements outlined above identify the sociometric measure in a more or less pure form, and are generally in agreement with Moreno's definition. However, relatively few studies in this area, perhaps not more than a quarter, meet all these requirements. For other studies, especially for those that do not permit restructuring of the group or do not ask for choice in terms of specific criteria, the terms "near-sociometric" and "quasi-sociometric" have been used. We do not maintain this distinction here, although we emphasize the importance of reporting in detail all the relevant conditions of any study. There are many circumstances where modified sociometric techniques are the only ones that can be used, and others where modifications lead to more meaningful results.

One of the most frequent modifications is in the limitation upon the number of choices the individual may make or, more accurately, *the specification of the number of choices the individual is required to make.* This alteration has certain obvious disadvantages in that it reduces the range of interpersonal reactions that the measure assesses. Eng and French (1948) found that when they used a ranking derived from a paired comparison of each member in the group as a criterion, unlimited choices showed a closer correlation to this criterion than either three or five choices. In certain circumstances it is of considerable interest to examine the individual variability in number of choices the subjects make when there is no limit provided. Further, if we wish to distinguish between the unchosen person and the isolate, the isolate being the person who neither chooses nor is chosen, we need an unspecified number of choices. In spite of these advantages adhering to the unlimited choice, there are sufficient disadvantages in terms of subjects' time required, rapport, and the eventual statistical analysis of the data, so that under many circumstances it is desirable to specify the number of choices to be made. When choices are limited or required it is roughly conventional to specify three.

Frequently the technique is used without the promise of restructuring the group. However, even overlooking the social obligations of the investigator, it would appear wise to restructure the group, as this will increase the sensitivity or care with which subjects report choices and rejections. On the other hand, the impossibility of fulfilling such a requirement does not mean the data will have no value. In many cases, because of the nature of the group or the nature of the criterion, it is impossible to introduce such changes. Subjects frequently give careful and accurate judgments in the absence of immediate and personal benefit. Thus, we may view restructuring the group as desirable because of the added realism and motivation it provides for the subject but, at the same time, it is only one of many factors that determine the relation between subject and research worker and it should not be considered an indispensable requirement.

The customary precondition that the groups be small enough to permit face-to-face interaction is often altered to the needs of the study. Moreno's original study (1934) was of a closed community, but it should be noted that it was large in size. The more general use of the measures has been in smaller closed groups. However, there has been increasing need for and actual application of the techniques in large groups and open communities (Lundberg and Steele, 1938; I. A. Stewart, 1948; F. A. Stewart, 1946, 1947a,b). In these the problems of statistical analysis are greatly increased (Bassett, 1948; F. A. Stewart, 1948), but the fundamental application remains the same. Subjects are still asked to select individuals they would like to have near them or participate with in particular activities, although in this case it is not known what the group is from which the subject actually chooses—that is, we do not know exactly what the population of experienced individuals is for the subject. Similarly, the customary precondition that the groups have existed over a period of time is often altered. Barker's (1942) investigation of "first impressions" is an example of a study where it was necessary to administer the sociometric techniques *before* any prolonged interaction.

SOME RELATED DEVICES

There are an appreciable number of techniques that fail to meet the definition proposed above but still remain sufficiently similar to the sociometric test to deserve consideration.

Sociometric self-rating or relational analysis. This technique is an extension of the sociometric test and is used in conjunction with it. It involves the use of the sociometric self-rating or prediction of sociometric choice, e.g., "Who do you think will choose you?" This method was first mentioned by Moreno (1942) in connection with the training of subjects' perception of others through repeated sociometric trials. Maucorps (1949) used a restricted form of this technique in a study of French Army officers, as did Lundberg and Dickson (1952) in their study of interethnic relations in a high school population. These investigators limited the predictions to those individuals whom the subject had himself chosen. A similar application was made by Williams and Leavitt (1947) in a study of Marine Corps officer candidates. Tagiuri (1951, 1952) has elaborated and developed this approach under the title of "relational analysis," and has indicated its usefulness in the study of social perception. There seems little doubt that the possibility of measuring the accuracy with which the individual perceives the choices and rejections that he secures from his associates provides an important additional dimension that may be expected to have significant psychological correlates.

Group participation scale. Pepinsky, Siegel, and Vanatta (1952) used an approximation of the Thurstone Method of Equal Intervals in an attempt to develop a sociometric index that would measure "effectiveness of participation in group activity." One hundred and seventy sociometric items were submitted to an undergraduate population for judging, and subsequently a reduced list of 40 items was submitted to a group of professionally trained psychologists. At the end of this process it was possible to select a group of 24 items that represented each of eight scale positions with three items all of which showed a reasonably small amount of variation among judges. In summarizing the results of their study they state: "The Group Participation Scale, used as a sociometric test in four fraternity groups, yielded (1) almost the entire range of scores as defined operationally by our scale, (2) a more normalized distribution of scores than seems to be provided by most sociometric tests, and (3) a relatively high agreement . . . of group members in choosing given individuals" (pp. 418-419).

Group preference record. This variation of the sociometric technique requires that members of the group respond in terms of like, dislike, or indifference to every member of the group (Newstetter, Feldstein, and Newcomb, 1938; Zeleny, 1939, 1940a). Thus, it differs from the conventional sociometric test primarily in that it forces an explicit evaluation of *all* group members by the subject. It has frequently been used in combination with some system of weighting choices.

Multirelational sociometric survey. An extension of the customary sociometric test by systematically combining it with the collection of additional data has been suggested in several recent publications (Weschler, Tannenbaum, and Talbot, 1952; Massarik, Tannenbaum, Kahane, and Weschler, 1953). In this adaptation two classes of questions (criteria) are utilized, i.e., organizational "goal directed" and "non goal directed." For each goal-directed activity five relations are investigated: (1) the prescribed, defined by leaders of the organization or the table of organization; (2) the perceived, defined by the subjects' reports of the prescribed relations; (3) the actual relations which the subjects report; (4) the desired, relations normally ascertained by a positive sociometric test; and (5) the rejected, the relations normally ascertained by a negative sociometric test. With these five sets of data, expandable by the number of criteria used, it is possible to establish a large number of meaningful indices for the classification of relational patterns in an organization.

Estimates of time. An alternative procedure to the sociometric test was suggested by Moreno, Jennings, and Sargent (1940) in the use of estimates of time as a means of measuring the intensity of choice. In this procedure the investigator asks the members of the group to estimate the proportion of time, given a finite limit, they would like to spend interacting with the other members of the group in a given activity. This technique has been further adapted by other researchers (Stogdill and Shartle, 1948; Stogdill, 1949) to require that the subjects estimate the actual amount of time spent in contact with various group members. A further modification involves securing information about actual visiting patterns in a community setting (Loomis and Davidson, 1939a,b).

The Moreno and Jennings modification remains closely related to the conventional technique, as the subjects are still requested to make a choice at the level of wish or desire. The use of time is primarily to permit the weighting of choices made. The other modifications, however, take the response out of the level of wish and ask for a factual report of interaction. These are thus perhaps closer to customary observational techniques than to a sociometric test.

Guess who technique. Another technique related to the sociometric test is the use of identification or "guess who" questions. Actually the origins of this method are as early as those of the sociometric technique, since they go back at least to the character studies of Hartshorne, May, and Maller (1929). This technique involves presenting the subjects with various behavior descriptions and asking them to "guess who" among the members of their group this description best fits. By varying the descriptions along various kinds of positive and negative continua it is possible to get, in terms of the frequency with which they are selected for the various descriptions, an ordering of the subjects that bears rough correspondence to sociometric data. In many cases this kind of data has been used to supplement conventional sociometric techniques and provide a picture of the reputations or impressions created by the subjects on other members of the group (i.e., the manifest personality). This technique may possess some advantage over the sociometric test in that it is slightly more indirect and the intent of the investigator is less apparent to the subjects. This would be an advantage only where there would be reason to expect that subjects would attempt to interfere with or disrupt the sociometrists' intent. Moreover, the "guess who" device is largely limited to "choice" or "reputation" scores and does not provide much information about patterns of interaction or communication.

As we have said, sociometric techniques merge into general methods of systematic observation. No attempt will be made to discuss observational techniques here, as the reader may refer to Chapter 10 for a complete account of such methods.

Analysis and Representation of Sociometric Data

The general increase in quantitative sophistication of social science investigators is nowhere better reflected than in the use of sociometric measures. Procedures and statistical techniques today are strikingly differentiated from those in common use some twenty years ago. During this period there has been a transition from descriptive, nonstatistical procedures of analysis to a state where virtually all investigators concern themselves with quantifying their data and most make some attempt at statistical analysis. Our discussion will consider techniques for the analysis and representation of data under the following headings: graphic methods, simple quantitative methods, statistical methods, matrix methods, and fractionation of groups.

GRAPHIC METHODS

In addition to introducing and popularizing the sociometric technique as a means of collecting valuable empirical data, Moreno also presented a method for summarizing the results of this technique that has proven most influential. This method, the *sociogram,* was first described in *Who Shall Survive?* (1934) and during the following decade or more served as the primary method of presenting sociometric data.

The sociogram is a diagrammatic device for summarizing the choices and rejections among members of a group. It employs geometric figures to represent members of the group and various kinds of lines joining the figures to represent choices and rejections. At this point there is no single convention for the drawing of diagrams but, rather, there are many alternatives available to the investigator.

As a consequence of continued use of sociograms in a variety of different settings, a number of patterns of relationship that are frequently encountered have been described. The ultimate number of such patterns is virtually limitless. Considerable attention is paid to types of sociometric patterns by Moreno (1934). Here we shall consider briefly a few of the patterns of psychological interest that are commonly seen.

The *overchosen* is the member of the group who receives a disproportionately large number of choices from other members of the group. Moreno (1934) has used the term "star" to refer to all varieties of popularity and leadership, whether or not they involve a large number of individuals centering their choices around the star. As the iconic qualities of the term seem to imply one figure surrounded by a number of admirers (Chapin, 1950; Moreno, 1950a), it is perhaps best to reserve it for that specialized form of popularity or personal choice where one individual is chosen by a number of others who do not direct a large number of choices among themselves. The term *overchosen* seems appropriate as a general term to refer to the individual who receives a large number of choices no matter what the pattern is in which they occur. Overchosen persons may be leaders in the group, but this is not necessarily the case, as has been indicated frequently (Chapin, 1950, 1951; Criswell and Jennings, 1951; Jennings, 1947, 1950; Moreno, 1934, 1950). It is evident that an individual who is not overchosen may occupy a position of leadership by virtue of receiving certain key choices. A term such as *power figure* may be appropriate for referring to such a person.

The *isolate* is the individual who receives no choices and makes no choices (or receives no choices and rejections and makes none). The isolated person is essentially not in the group; other members are either indifferent to him or do not respond to him, and he does not display attraction to or rejection of the other members. In the literature the *unchosen* and the *underchosen* have often been called isolates.

The term *rejected* has ordinarily been reserved for persons who receive a large number of rejections.

The *mutual pair* has received considerable attention, and the number of mutual pairs in the group has ben used as the basis of several sociometric indices. The term *triangle* is usually reserved for the case where three individuals are connected by mutual choices. Larger units of persons connected by mutual choices are usually called *chains*, and closed chains or rings are possible, as well as many other variations, according to the bonds involved.

Since the variety of patterns which occurs is so great, the above terms have frequently been applied without precision. It is obvious that the researcher should specify the exact meaning that he gives the terms he uses in his descriptions and analyses.

The sociogram has received criticism from a number of quarters because there have been many examples of "unreadable" sociograms. In many cases the fault lies in poor drawings, but more often the trouble has been that the persons drawing the sociograms have not had as their purpose the analysis of data, but rather the presentation of a persuasive visual device. Several investigators have made attempts to standardize the drawing of sociograms and thereby remove a part of their subjectivity and unreliability, as well as making the diagrams more readable. Some of these efforts have attempted to introduce elements of quantification.

Proctor and Loomis (1951) have suggested a set of conventions to be applied in determining the distance separating individuals in sociogram positions. Essentially, they develop an index of social distance based upon the reciprocal relation existing between each pair of individuals in the group. They then propose that each point along this continum of social distance can be equated to some arbitrarily designated distance on the sociogram. These investigators assume that if x chooses y he is closer to y than he is to other individuals whom he did not choose. Further, if x rejects y he is further away from y than other individuals whom he did not reject. Finally, the combination of mutual choice or mutual rejection is viewed as respectively closer and further away than the single choice or rejection. If we overlook the symmetrical duplications that are possible, e.g., x choosing y and y rejecting x, as opposed to x rejecting y and y choosing x, there are only six different kinds of interpersonal relationship. It is possible to overlook the symmetrical variations, as each variation implies the same distance or closeness. The types of interpersonal relations are:

Type A: i chooses j, and j chooses i
Type B: i chooses j, and j ignores i
Type C: i chooses j, and j rejects i
Type D: i ignores j, and j ignores i
Type E: i ignores j, and j rejects i
Type F: i rejects j, and j rejects i

Making the further assumptions that types D and C are equidistant, Proctor and Loomis propose that it is now possible to array individual pairs on a continuum from close to distant, or choice to rejection, with type A representing the closest relationship and type F the most distant or aversive. Actually, if one objects to the assumption that a choice and rejection are equivalent to two "ignores," there is no reason why these two types could not be made separate points on the distance continuum, perhaps with the choice and rejection being placed closer than the mutual "ignorers." The major intent of a sociogram drawn in this way is to cluster closely related persons and thus to make it easy to identify cliques. There is some danger, however, that this forced concentration may actually mask networks of communication, especially if the group is large and there is a high incidence of choices.

This procedure provides the investigator with a definite goal toward which he may aim, even though it will not always be attainable. At the same time it introduces a certain standardization into the practices of different individuals constructing sociograms. It is important to realize that the larger the group the less likely it is that the conventions of this system can be satisfied. When a subject is placed in relation to a large number of individuals who are simultaneously in relation to other individuals, and independent demands are made as to the distance separating all these positions (persons), these demands will frequently not be satisfiable on a single plane surface.

Another attempt at standardization was made by Northway (1940) when she introduced the "target sociogram." In this diagram a series of concentric circles are used to designate areas in which individuals are to be placed according to their choice status. Usually, the field is divided by quartiles, and the center of the target is the location of the individuals in the quarter receiving most choices, the three bands around the center receiving the persons in the respective quarters. Thus, the target sociogram identifies the relations among the group members in the usual sense, and further identifies the choice status of the individual at the same time. In this form the target diagram may represent only a slight improvement over the ordinary sociogram in identifying individuals who are highly chosen, since in the sociogram such individuals may be identified by their spoked-wheel appearance on the diagram. However, if it is important to locate all highly chosen individuals at one time, this form of the target diagram is quicker to read. Northway and Quarrington (1946) further suggested the use of sectors to separate members of the group in the target sociogram by characteristics such as sex or ethnic background, in order to facilitate examining the data for cleavages. Bronfenbrenner (1944b) proposed the possibility of using the target sociogram with the concentric areas designated according to the probability with which the number of choices received exceed or are less than the expected number.

The Northway target diagram may be viewed as useful in indicating highly chosen and relatively unchosen persons, and quantification is relatively explicit, but this presentation may not be intrinsically superior to other varieties of the sociogram, e.g., one which uses identification symbols of different sizes to designate the relative number of choices received by the person. An example of the latter is seen in the sociograms used by Powell (1951). Further, the use of sectors in a target diagram to designate sex or ethnic background distinctions can be duplicated in ordinary sociograms, e.g., Loomis, Beegle, and Longmore (1947).

Borgatta (1951) suggested that the primary principle in sociogramming should be that of minimizing the number of crossed lines. Sociogramming according to this procedure does not restrict the number of persons that can reasonably be handled as severely as some techniques. It has the further advantage that the minimum number of crosses may be achieved only in a limited number of ways, and the alternate sociograms are equally readable.

The kind of sociogram used is limited only by the investigator's needs and imagination. There are many alternative methods of presentation of sociometric data, and these may even include three-dimensional sociograms (Chapin, 1950; Weschler *et al.*, 1952). The sociogram does not restrict the investigator to one variable, as multiple variable sociograms may be handled by using lines of different colors, overlays, or other devices.

An appraisal of the sociogram must begin with a heavy credit to the method for its dramatic and compelling qualities. The possibility of summarizing group relations in terms of a series of graphic conventions that appear directly understandable and interesting to most audiences is an important asset. Particularly in circumstances where the information is being presented in an "action" context, the possibility of easily engaging an audience without the necessity of presenting technical information or complex explanations may be of crucial importance. In many industrial and military applications, as well as in some educational settings, the "glamour" of the sociogram makes it almost a prerequisite no matter what additional techniques may be used in analyzing and presenting sociometric results. For small groups the difficulty of constructing sociograms is slight compared with the labor involved in some of the other analytic techniques to be discussed. A final advantage of the sociogram is that it presents results in context. The relations existing in the entire group are presented at once rather than being abstracted into a number of indices or scores.

In spite of these desirable qualities, there are negative attributes that inhere to the customary use of the sociogram. The representativeness of the sociogram has depended almost completely upon the vigilance and sensitivity of the investigator, with few or no external checks upon the adequacy of the diagram he presents. Thus, depending upon the particular spatial conventions or groupings the investigator uses, he can create different impressions of the group being examined. This absence of external or formal checks upon the procedure has led generally to relatively low inter-judge agreement using this technique. The suggestions of Northway, Proctor and Loomis, and Borgatta can be seen as important steps in the direction of reducing or eliminating this particular shortcoming. Where the group is very large, especially if unlimited choices and rejections are permitted, construction of a sociogram may not be feasible.

SIMPLE QUANTITATIVE METHODS

Simultaneously with the development of graphic techniques, there have been important developments in the use of quantitative techniques for analyzing and summarizing sociometric data. We shall begin our discussion of these developments by considering tabular methods, scores or indices, and profiles. In later sections we shall consider statistical and matrix methods.

Tabular methods. Sociometric data are summarized completely in the sociogram, but an equally complete and natural way of summarization is the $N \times N$ table or matrix, in which N refers to the number of subjects. The subjects are placed along the left margin of the table in some arbitrary order, and then the order is reproduced along the top margin of the table. The responses of a given subject to the remaining $N - 1$ other subjects may then be placed in the proper cells. Ordinarily the main diagonal of the table or matrix is empty, since subjects are not asked to respond to themselves. This technique makes the summarization of the number of choices and rejections made and received by each subject simply the task of adding across a column or down a row. It serves as the basis for summarizing data into simple scores, and then by various operations into particular indices, and also provides the matrix of raw data for the various matrix analyses.

Scores and derived indices. Quite early in the use of sociometric measures some investigators, e.g., Moreno (1934), chose to summarize what appeared to be important aspects of the choices and rejections within a group in terms of simple scores or ratios. These methods have frequently been taken to task as oversimplifying the complexity of the group relationships. In cases where the score has been reasonably derived so as to get at an aspect of the choice and rejection pattern of general interest, this approach has led to fruitful results. Thus, although the individuals who hold to an organismic or holistic view of the world may mourn the loss of the complex interrelations that are to some extent present in the sociogram, or the $N \times N$ table, the use of indices and their derivatives has become increasingly popular.

There are many ways in which scores and indices can be categorized. Some deal with or represent the individual, while others represent the group as a whole. Still others represent the individual in interaction with other members of the group. Our intention here is not to present a systematic summarization of indices, but rather to illustrate the breadth and variety of indices that have been used by different investigators.

Moreno's *Who Shall Survive?* includes the explicit formulation of a number of indices and implicitly assumes a larger number. The identification of the star, the isolate, and the rejected essentially provides the basis for a large part of the individual indices later developed. Moreno also proposed a number of group indices: *emotional expansiveness* refers to the amount of utilization of available choices. This index can be used only where the number of choices is not required, i.e., the subjects are permitted to make unlimited choices or as many as they wish up to a certain limit. Although the index is customarily used by Moreno in the description of groups, he also uses the concept in the individual sense. The *ratio of attraction* is equal to the number of attractions divided by the total number of attractions and rejections, converted to a percentage. The *ratio of repulsions* is similarly computed. Moreno set the ground work for an index of "cleavage" in a group in his discussion of restriction of spheres of choice, and more directly in the use of the *ratio of interest for home group* and for *outside groups,* which are computed on the basis of available choices utilized. The *index of relative popularity* (of a group) refers to the proportion actually received of the possible choices the group could receive from the external groups.

The number of choices received serves as the basis for a number of indices. Individuals may be classified arbitrarily, e.g., they may be stratified by quartiles, deciles, or other dividing points. An example of this is Jennings' (1950) stratification of persons in terms of dividing points established on the base of the mean number of choices received. She identifies one standard deviation above the mean number of choices received as *overchosen,* and those one standard deviation below the mean as *underchosen.* Arbitrary classification (or fractionization) of this type extends the utility of the data in allowing the possibility of comparison to or combination with data from additional groups, since the size of the group is considered in the classification. However, it should be noted that being one standard deviation above the mean in one group may have different meaning than being one standard deviation above the mean in another group. Still another classification is presented by Proctor and Loomis (1951) in their definition of *choice status* as the number of choices received divided by one less than the number of persons in the group, with a similar definition for *rejection status*. An important social status index has recently been reported by Katz (1953). This measure employs vector concepts in order to take into consideration not only the number of choices but also the social position of those who choose.

The sum of acceptances plus the sum of rejections, divided by one less than the group size, is an index which McKinney (1948) identifies as *social status index*. Smucker (1949) identifies this index as *social intensity*. *Choice rejection status* is identified by Proctor and Loomis (1951) as the number of choices less the number of rejections received. Such an index, however, may assume negative values, and should be adjusted where necessary by addition of a constant factor if only positive values are desired.

The number of mutual choices in a group serves as the basis of several indices. McKinney (1948) identifies a *compatibility index* as the number of mutual choices divided by one less than the group size. Proctor and Loomis (1951) identify *group cohesion* as the number of mutual pairs divided by the possible number of mutual pairs in the group (as restricted by instructions for choosing). The same authors (1951) identify *group integration* as one divided by the number of persons receiving no choices.

The number of indices introduced to handle sociometric data is so vast that it is nearly impossible to enumerate them all and list the various dimensions with which they have been identified. Proctor and Loomis (1951) present an outline of indices that may serve as a beginning point for persons who wish to examine a range of the defined indices in sociometry. In addition to those we have already mentioned, Proctor and Loomis review such indices as *in-group preference, in-group cleavage, in-group cohesion,* and *in-group climate.*

As the reader can see, the sociometric investigator faces many possibilities for devising indices. We have already indicated that persons may be ranked within a group on the basis of the raw number of choices made and received, rejections made and received, and combinations of these. Such indices are of immediate use where the sociometric status of individuals in a given group is to be examined in relation to other criteria. The problem of comparability of persons selected from two or more groups may be handled in several ways. Assuming that choices and rejections have been made under the same directions, persons might be ordered by their relative positions in the group, as with Northway's quartiles or Jennings' standard deviation classifications. In such cases the investigator should be aware that groups vary in the distribution (and range) of choices made, and if he is willing to make certain assumptions, he is able to transform his data so as to increase its utility. An index based on the number of persons in the group will not necessarily be much of an improvement over the use of raw scores, since this merely divides all cases by a constant, and will not increase the comparability of persons in groups that differ greatly in size. Rather, in cases where persons are selected from a group of say 20 persons, and others from a group of say 200, comparability on the basis of such an index would probably be more questionable than the use of raw scores as an index. Essentially, the problem of what type of index is proper in the situation is defined by the use to which the index is to be put, and further, the types of assumptions the investigator is willing to make.

We have not yet discussed the problem of intensity of responses, that is, the differential meaning of first, second, third, etc., choices. Some investigators (Zeleny, 1939, 1947b) make extensive use of ranked choices in their analyses and incorporate them into their indices. This simply requires an adjustment in the computation of the index, and researchers may adapt indices in this way without serious difficulty. However, very little work has been done at a theoretical level on the meaning of intensity in sociometric choices (Foa, 1950) or the value of assigning arbitrary weights. The researcher should consequently be cautious in weighting choices, especially if the choices are required. We shall discuss later Moreno's (1934) report on the greater stability of first as against second choices which might lend support to the weighting of choices, but since weighting is ordinarily arbitrary, there is no evidence that equal weights (ignoring choice order) would not be as good as assigning weights by any other arbitrary technique.

Similarly, we have not dealt with indices derived from choices made on several criteria. Converting these data to customary indices is a relatively simple matter, but as we shall indicate in our discussion of reliability, the investigator should be aware of the assumptions implicit in his operations. One method of dealing with responses to several criteria is through the use of profiles.

Profiles. The investigator may wish to examine the choice status of the individual in terms of a number of indices or criteria at the same time. In this case, some unit of standardization is necessary among the scores. Jennings (1943, 1950) has dealt with profiles of individuals based on (1) choices made, (2) choices received, (3) positive reciprocations, (4) rejections made, (5) rejections received, and (6) negative reciprocations. The first three are given as the *choice pattern* of the individual, the last three as the *rejection pattern*, and the six together as the *choice-and-rejection pattern*. The pattern is identified by the ordered series of six pluses or minuses, a plus signifying that the person is above the mean in the particular item and a minus signifying that he is below the mean. Such presentation allows the researcher to examine the distribution of patterns and, essentially, the relationship between performances on the individual items, by examining the distribution of types. Profiles are reported elsewhere (McKinney, 1948, Smucker, 1949) in various contexts, but the contribution here is in standardization on the same criterion to facilitate examination of types. Bjerstedt (1952) extends the work of Jennings and provides a systematic method of classification which is easily extended to at least two forms of sociographing. Bjerstedt's innovation is especially interesting in the amount of information which is made available for immediate reading, as well as the fact that sociographing according to his specifications is reproducible by independent investigators.

STATISTICAL METHODS

Statistical techniques are widely used in the analyses of sociometric data both to test the significance of observed findings and also to provide derived scores or indices. Several investigators who have used quantitative methods have developed scores with the intent of distinguishing between differences or deviations that should properly be assigned to chance factors as opposed to those that represent actual, replicable deviations from the group norms.

Bronfenbrenner (1943, 1944a) developed the notion of the chance model, introducing the Pearson Type III Curve as the most adequate approximation to the binomial expansion for sociometric data, and elaborated on the concept of statistical significance as applied to sociometric data. Distribution of indices on a chance basis, given a level of significance, provides the upper and lower limits of variation beyond which phenomena may be judged to be statistically different from chance. Thus, for example, given any distribution of choices received in actual data, using the limits established by the chance model, persons may be identified who receive significantly more or less choices than would be expected by chance. Bronfenbrenner's approach is intended to have the advantage that it provides a constant frame of reference for comparison or pooling of data; that is, for example, persons who receive significantly more choices than would be expected by chance may be recruited from several different groups and pooled for comparison with a similarly selected group of underchosen persons. Or, persons who receive significantly more choices than would be expected by chance in one type of group may be compared with similarly selected persons in another type of group.

The usefulness of a probability model is unquestioned, although there are many discussions in sociometric literature of misapplications or the adoption of incorrect (or incomplete) assumptions (Criswell, 1947; Proctor and Loomis, 1951). One example of an inappropriate assumption is Bronfenbrenner's (1943, 1944a) suggestion that the distribution of mutual choices will be according to the binomial expansion. Criswell (1946, 1950), Edwards (1948), and Loomis and Pepinsky (1948) have all indicated the theoretical inadequacy of this formulation. Criswell's (1939) assumption that chi-square values may be ordered rather than treated as significant or not significant has been objected to by Loomis and Pepinsky (1948).

One special consideration in the use of Bronfenbrenner's method should be noted. Separating persons as significantly overchosen and significantly underchosen provides an aura of respectability in the handling of the data, but such a procedure may be unrealistic. Such classes are frequently very small and require a large number of samples, or large samples, to permit further statistical comparisons. On the other hand, there is good evidence that the stability of a person's choice status is high, so that ordering of persons on choice status will have reasonably constant meaning. In using only classes of persons who are significantly above or below chance expectation, the researcher appears to make the assumption that those who do not satisfy this condition of statistical significance receive their choice position on a basis of chance. This assumption does not seem necessary and results in the loss of large amounts of data. Thus, unless the investigator is genuinely interested in the extremes of the distribution, there seems good reason for him to use in comparisons subjects who are selected arbitrarily or in terms of some rational criterion rather than only those who deviate from the mean at some particular level of significance. Criswell (1950) has posed this criticism in a slightly different form in her detailed analysis of the constant frame of reference problem.

In contrast to Bronfenbrenner's method, which uses probability values as scores, Criswell's scoring procedures attempt to eliminate the influence of chance distributions. The first scoring method suggested by her (1939b) was for the measurement of race preference expressed by a group as a whole. She devised a ratio in which a given racial group's preference for itself over another racial group was expressed as:

$$\text{Group self-preference} = \frac{SO^1}{OS^1},$$

in which S = number of choices given by a group to itself, S^1 = number of choices it would give itself by chance, O = the number of choices the group gives to the other racial group in the experimental population, O^1 = the number of choices it would give by chance to the other racial group. Another method of probability scoring devised by Criswell (1946, 1947) defines group coherence as the obtained ratio of reciprocated to unreciprocated choice divided by the corresponding chance distribution. The formula is then:

$$\text{Group coherence} = \frac{\text{Number of reciprocated choices} \times q}{\text{Number of unreciprocated choices} \times p},$$

where p/q is the chance ratio of reciprocated to unreciprocated choices.

Statistical methods have also been used to examine the likelihood that a given finding is the result of chance factors. The first use of the chance model with sociometric data was made by Moreno and Jennings (1938) when they set up an empirical random selection procedure and generated sociometric groups. They then compared these groups in terms of the distribution of choices with a probability model (provided by Lazarsfeld) based on the binomial expansion and found no apparent differences. The chance model was then compared with actual data by means of the chi-square test, and significant differences between the obtained and chance distributions were immediately apparent. In the actual data, unchosen and highly chosen persons occurred more frequently, and mutual pairs also occurred more frequently than in the chance distribution.

Several attempts have been made to apply statistical methods in testing differential choice patterns within a group. The first of these was suggested by Criswell (1939), who used a chi-square test in which:

$$\text{Chi-square} = \frac{(\text{Self-preference expressed} - \text{Theoretical self-preference})^2}{\text{Theoretical self-preference}}.$$

A critical discussion and extension of this approach is provided by Loomis and Pepinsky (1948). Seeman (1946) presents an alternate procedure in his study of intragroup Negro attitudes, in which he uses an analysis of variance design. It should be noted that Bronfenbrenner's (1943, 1944a) approach, using the Pearson Type III curve, may also be employed here.

MATRIX APPROACHES

In recent years there has been a rapidly accelerating interest in the analysis of sociometric data through methods of manipulating the $N \times N$ table on which sociometric responses are summarized. These methods of matrix analysis are still in the process of development, but for certain kinds of problems they show marked advantages over other methods. In particular, they are explicit and repeatable, and lend themselves to statistical analysis; some of them deal with large amounts of data relatively efficiently, and certain graphic features are retained.

The first of the matrix approaches to sociometric data was developed by Forsyth and Katz (1946). In their method, the $N \times N$ table is taken as the raw matrix. The original $N \times N$ matrix is then changed through moving pairs of rows and columns (the order of columns must always remain identical to that of the rows) to reveal subgroups within the structure, stars, isolates, and other characteristic forms. The ordering is one of trial and error, in large part, with the main goal of forcing mutual choices to gravitate towards each other or, more specifically, to concentrate choices along the main diagonal. Presumably, if rejections are included in the data they will gravitate away from the main diagonal as the choices gravitate towards it. The recommended procedure is to pick two persons with mutual choices and place them adjacent to each other in the matrix. Then select another person who has mutual choices with both, if such a person exists, or who is related by choice with one or both of the original two. Rearrange rows and columns as is necessary to maximize the concentration. When no persons can be added to this nucleus of individuals, consider this a subgroup and begin the rearrangement of the remaining individuals. Continue the procedure until all persons are related to a subgroup, or until only isolated or rejected individuals remain. Subgroups should be rearranged so that the maximum concentration of rejections (if any) is located in the corners of the matrix away from the main diagonal. Individuals who choose into a subgroup but are rejected or ignored are placed on the upper right boundary of the relevant subgroup, and persons who are relatively underchosen but apparently related to the subgroup may be placed at the lower left boundary of the subgroup; where both conditions are apparent, the second takes precedence in the arbitrary placement. Genuine isolates should appear near the center of the matrix. At this point the matrix may be re-examined to emphasize the concentration of subgroups, etc.

The advantages of this matrix approach as suggested by the authors include the following: Different investigators will tend to produce the same or similar matrices. Subgroup structures are evident in the concentrations along the main diagonal. Potential conflict betwen two subgroups may be indicated by concentration of rejections in the two intersections of the subgroups, and the relative rejection of one subgroup for another may similarly be examined. Relative permanence of the group structure may be indicated by the extent to which subgroups are "well knit" and have absorbed satellites. Isolates and the overchosen may be identified as before by column and row totals and, in addition, one may see whether the overchosen are chosen primarily by one subgroup, or two, or from the group at large, and similarly, whether a rejected person is rejected by a group in particular or from the group as a whole. The major disadvantage of the technique is that the manipulation of rows and columns is cumbersome.

Katz (1947) indicated as additional advantages of the matrix procedure that not only could positive and negative choices be included in the matrix but that, essentially, indifference is accounted for, and there is no necessary restriction that excludes the utilization of weighted choices. Moreover, choices may be unrestricted, limited, or required. The major contribution of this second paper, however, is in describing a canonical form for the matrix which would serve as the preferred ordering, namely, the maximization of concentration of positive choice about the main diagonal. Essentially, the form may be stated as follows: Given N row vectors and N column vectors, and the interval from one row (or column) to the adjacent one taken as one, the square of the distance of the element in the intersection of the ith row and jth column from the main diagonal is $\frac{1}{2}(i - j)^2$. Thus, ignoring the constant, the least square solution of this form is defined by the minimum value of $\Sigma_i \Sigma_j e_{ij} (i - j)^2$, where e_{ij} is the element of the ith row and jth column. Once in canonical form, the matrix is subject to various statistical operations. However, because of the difficulty of bringing a matrix to canonical form, these statistical suggestions of Katz are primarily of academic interest at present.

The utility of the matrix approach suggested by Forsyth and Katz is quite obvious, but the limitations should not be underestimated. Moreno (1946) commented on the presentation of Forsyth and Katz and suggested that he viewed the matrix approach as supplementing the sociographic approach rather than replacing it. He indicated that certain types of relationships, such as chains of communication, are quite difficult to locate in the matrix, while they are quite simple to trace on the sociogram. The "cumbersome" aspect of the matrix manipulation should not be underestimated. The fact that a canonical form has been described does not make the task of manipulating the matrix any easier; it still remains a matter of sophisticated trial and error to get as close to the canonical form as possible. In some cases this may be as time consuming as drawing and redrawing a number of sociograms. On the other hand, the Forsyth and Katz approach may have an advantage when the number of choices and rejections made by the subjects is very large. In sociograms of the Loomis variety the problem of congestion would be evident, and in sociograms minimizing the number of crossing lines the number of such lines would be very great. In such circumstances the Forsyth and Katz matrix should make the subgroup structures more evident. A final objection to this method is that it forces the assumption of linearity upon the data examined when it may be that this assumption is not warranted.

A matrix approach related to that of Forsyth and Katz is presented by Beum and Brundage (1950). Their method is one which is designed to maximize choices about the main diagonal. The method makes no restrictions on the data to be entered in the cells. The procedure is based on the demonstrable condition that:

> If weights are assigned to the rows of a sociomatrix (from 1 to N beginning with the bottom row) . . . and the average product of the elements in each column and the corresponding

> weights is maximized for each column, the sum of the squares of the elements about the principal diagonal is minimized. (p. 142)

This would produce a canonical form defined in a way similar to that of Katz (1947). The procedure for achieving the canonical form is not one of trial and error, as in the Forsyth and Katz approach, but one of a systematic iterative process. When iteration of the procedure no longer produces changes in the structure (arrangement of rows and columns), the canonical form is reached. When iteration produces recurrence of arrangements of rows and columns, alternate and equally "good" arrangements are displayed. The procedure has the advantage over the Forsyth and Katz method that it is more systematic (different investigators get the same results). However, the authors point out that the iterative procedure converges slowly, and that consequently the labor involved may be excessive.

Bock and Husain (1950) consider the Forsyth and Katz (1946) matrix approach similar to identifying "clusters" of psychological tests. They introduce the notion of ranking choices in terms of the "relatedness," that is, the degree of reciprocation (this assumes the subjects ranked their choices). One generates a new matrix by replacing the original values by the rankings in this dimension. Having made this conversion, the matrix may be analyzed using Holzinger's *B*-coefficient. The procedure is not unlike that of the Forsyth and Katz approach of beginning with two persons who are mutually related, and then picking additional persons who are highly related to the initial nucleus. If the Bock and Husain technique offers any single advantage, it is that a reasonable but quite arbitrary decision point can often be established to determine whether the last person added to the subgroup actually belongs to the subgroup, but it is not certain that this is actually better than what could be accomplished by inspection.

Luce and Perry (1949) with Festinger (1949) have developed a second method for matrix analysis of sociometric data which is based on matrix multiplication. This method is immediately differentiated from the Forsyth and Katz approach, as it places severe restrictions on the type of data which can be handled. Only choices *or* rejections on a single criterion can be treated and there can be no weighting of responses.

The $N \times N$ table or matrix is again the beginning point of this method. The basic operation involved is raising the raw matrix to various powers. The simplest problem for which the method may be used is to identify chains of communication, and the easiest of these to find is that of the two-step chain, where individual *i* chooses individual *j* who in turn chooses someone else. Given an $N \times N$ matrix, the two-step chains can be identified through squaring the matrix. The operations involved in raising the matrix to various powers are described by Luce and Perry (1949), Festinger (1949), and Festinger, Schachter, and Back (1950). Given the squared matrix, the entry in the *ij*th cell indicates the number of two-step paths from *i* to *j*. The cells on the main diagonal contain the number of mutual choices in which each individual is involved. The total number of mutual choices in the group is obtained by summing the cells of the main diagonal and dividing by two, as each mutual choice appears twice. Through raising the matrix to higher powers, cubing, etc., it is possible to use the method to identify cliques. Ordinarily this method has defined cliques in a highly restrictive fashion, asking that each member of the subgroup select all other members of the subgroup. Luce (1950), however, extended the generality of this approach through indicating how it is possible to alter the definition of the clique to include looser structures, i.e., identifying cliques in which there were indirect or incomplete patterns of choice among the members.

It is important to realize that having identified particular kinds of patterns or structures it is virtually always necessary to return to the original matrix in order to identify the persons who make up the pattern. One should also be aware of the fact that as the matrix is raised to higher powers *redundancies* will occur, that is, the same pattern may be counted more than once. A method for determining redundancies in five- and six-step chains is proposed by Ross and Harary (1952).

The use of matrix multiplication in the analysis of sociometric data has the greatest potential for the development of standardized techniques for the identification of structures and channels of communication. However, the nonmathematically inclined may find some difficulty in understanding the procedure. At present the techniques are not sufficiently well developed for proper assessment. Nevertheless, the following difficulties are evident: (1) The limitation to unweighted choices *or* rejections may be found restrictive. (2) The matrix operations are laborious unless done on IBM equipment, and programming for this matrix procedure may be more effort than it is worth. (3) Even as extended by Luce, the definitions of structures which may be isolated are restrictive. Identification of patterns of a given definition will still leave many other patterns unidentified, requiring either retreatment or neglect of the data. (4) The holistic view possible with the Forsyth and Katz method and the sociogram approaches is not possible. In favor of the method, it is the only approach available which is relatively unrestricted as to the size of the group which may be examined, and this alone may be sufficient reason for giving it attention. In addition, treatment of the data is readily repeatable in the hands of different investigators. Persons interested in reading relatively nonmathematical expositions of this approach should refer to Festinger (1949) and Chabot (1950).

Machine computation. It is obvious that some of the analyses we have discussed, the matrix multiplication as an example, if attempted by hand necessitate a prohibitive amount of labor. Fortunately, many of the operations required can be programmed and done on IBM equipment, and the investigator should become familiar with its use, especially if he is to deal with large samples. Beum and Criswell (1947) presented a method for machine tabulation of sociometric data. The technique has great versatility, and is unrestricted as to the size of the group to be considered. Most operations which can be reduced in form to simple equations can be done on IBM equipment either with the original cards or with decks generated in accordance with the programming of the operations.

Katz (1950) presents an alternate procedure for IBM tabulation. All operations possible by the Beum and Criswell technique are equally possible with this method, and for relatively small samples it may be more expedient. However, the technique has a limit of practical usefulness, since the number of persons in the sample is restricted by the number of available columns in the deck. Katz presents some sample procedures for getting mutual choices, measuring cleavages, raising a matrix to a power, trial and error manipulation for using the Forsyth and Katz method of matrix analysis, etc.

Persons planning analysis of large-scale sociometric data will do well to investigate the feasibility of using machine techniques, whatever the actual method of analysis. However, if the sample size is not large, the investigator may spend more time in preparing for analysis than other methods might require to complete the analysis.

FACTOR ANALYSIS

The possible application of factor analysis to sociometric data has been suggested many times, although as yet there have been only a few such extensions reported. Cervinka (1948) explicitly suggests the application of factor analysis to sociometric and similar structures in terms of isolating factors related to the manifest data, and Katz (1950), in his treatment of machine techniques, makes similar suggestions. Bock and Husain (1950) have carried out a factor analysis of a particular sociometric group and were able to identify two factors, "sex" and "reputation for science ability." Their presentation is primarily an exercise in which they demonstrate the possible utility of such an approach, especially in terms of factor scores computed for the group members. Proctor (1953) applied factor analysis to the classification of multiple sociometric questions.

FRACTIONATION OF GROUPS

The problem of identifying subgroups is basically a general one, and is encountered in most investigations. Much of our discussion of quantitative, statistical, and matrix methods has direct implications for this problem. Here we wish simply to stress again the wide range of possibilities for arriving at principles of classification and provide one example of a well-contrived solution to a fractionation problem.

Given a particular set of operations, the researcher may feel confident that he has the "best" way of dividing or ordering his sample into classes for further statistical treatment. Basically, however, the operations involved are arbitrary and should be accepted as such. While logic and theory are always important considerations there are points in the analysis of data where expediency must enter in to some degree. Thus, if less "exact" fractionation is used, the only question the researcher must ask himself is whether or not this lack of exactness could account for the failure to observe expected relations. Even if the classification is to be correlated with another measure, the "exactness" of the fractionation may not be nearly as important in the determination of the magnitude of the correlation as the number and size of the classes. Thus, as we have indicated before, statistical operations and quantitative respectability may not always lead to maximum utility. This point of view is consistent with Criswell's (1950) defense of the within-distribution fractionation, e.g., Jennings (1943), as against the use of classes of persons significantly above or below the chance expectancy.

At this point, let us examine an example of intelligent compromise in the process of fractionating groups. In considering the problem of differentiating status groups using multiple criteria (choices and rejections), Lemann and Solomon (1952), following Bronfenbrenner (1943), considered the possibility of making their high status group those individuals whose choice frequency exceeded the expected at the 1% level of significance and whose rejections fell below the expected at the 1% level of significance. The low status group would be the converse—falling below the expected in choices and exceeding it in rejections. The subjects that did not fall into these two groups would be placed in the middle status group. Partly as a consequence of the fact that the choice and rejection distributions did not coincide when one was inverted, they found that their logically defensible principle of fractionation was not workable. The vast majority of the subjects fell in the middle status group and there were few cases in either the high or low status groups. In the face of this experience they reasoned:

> . . . let the high status group consist of all those who received significantly more choices than would be expected, and who at the same time did not receive more than the mean number of rejections; and let the low status group consist of those who received more rejections than would be expected, and who at the same time did not receive more than the mean number of choices; and let all others not included in these two extreme groups make up the middle status group. (p. 31)

This method seems defensible, as it places in the high group all those who secure significantly more choices than would be expected by chance, unless they also receive a large number of rejections. The low status group is conversely defined, and the middle group is determined residually. Most important is the pragmatic outcome of reasonable numerical divisions of subjects into the three groups, and this was an important factor in the decision to employ this classification.

RELIABILITY

Discussion of the reliability of sociometric measures is fraught with the same difficulties that beset discussion of all instruments purporting to measure variables that change with the passage of time. In dealing with test consistency, at least, the sociometric tester finds himself sharing sentiments with the projective tester. Just as the user of the TAT or Rorschach protests, somewhat futilely, that he would be disastisfied with his test if it showed no change in an individual following the passage of time and concomitant changes in the individual, so the sociometric tester claims that his test must vary with the passage of time and changes in group structure.

The issue appears to be primarily one of the imputed effect of environmental forces upon the variable being measured. The psychologist's consideration of reliability has been heavily influenced by the definitions and procedures developed by "capacity testers" who for the most part were willing to assume that the variables of primary interest to them were not readily susceptible to environmental change. Such an assumption greatly simplifies the problem of examining instrument consistency, as it assigns all variability to test unreliability. For most psychological variables, however, it is empirically so thoroughly unwarranted as to make its adoption impossible. The user of sociometric measures seems bound to accept the probability of change over time in the variables he is interested in studying. Thus, the task of specifying the elements of change in test data that may be correctly attributed to the actual changes in the variable under study, as opposed to those elements of change that must be attributed to unwanted or chance fluctuation in the test, is a complex and exacting task.

Interpretive reliability. To begin with, it may be useful to distinguish between questions of reliability having to do with interpretation of the data provided by the test and the reliability of the data itself. Interpretive reliability is a much simpler problem and ordinarily is assumed in studies of test reliability. We are interested here in the extent to which the description of data by a given investigator, once the data are on hand, can be repeated by other investigators given the same data. Our discussion of the various kinds of analysis an investigator may employ has made clear that the interpretive reliability may be expected to vary with the particular technique chosen. Thus, if the data are analyzed in terms of an index the reliability of the procedure is apt to be very high. It is true that when the data are analyzed in terms of scores or indices the picture of the group that results may vary depending upon the particular quantitative choices the investigator makes. However, the problems are no different here than with any set of quantitative data that permits expression in terms of more than one set of scores. If the data are summarized in terms of a matrix and arranged according to procedures such as those suggested by Forsyth and Katz (1946), there should again be high agreement between two investigators treating the same data independently. It is only in the purely graphic representation that a great deal of unreliability is apt to occur (Loomis, 1948). Even here procedures such as those proposed by Proctor and Loomis (1951) and Borgatta (1951) reduce this variability considerably.

Thus, the problems posed by interpretive unreliability are not of great magnitude. Only when sociograms are used to analyze sociometric data is there apt to be excessive unreliability, and even this can be reduced by agreeing upon certain conventions in advance of analysis.

Test reliability. As we have implied, this is a more complex matter. In most situations it is not possible to provide definitive answers to the question of just how stable sociometric measures are. Better, one might say that it is very difficult to divorce the stability of the measure from the instability of the phenomenon being measured.

The particular problems encountered vary with the kind of consistency in which the investigator is interested. Thus, if we are concerned with the *repeat reliability* of the test, its capacity to produce consistent results over a period of time, appraisal of the reliability of the instrument is made difficult on the one hand by the effect of memory of original responses and on the other by the changes that the group may reasonably be expected to undergo. If we examine similarity of sociometric patterns over a short period of time we must account for a high degree of consistency not alone in terms of the consistency of the measurement but also in terms of the contribution of the subject's recall of his first test responses. On the other hand, if the two administrations are separated by a sufficient interval of time to make accurate recall of original choices unlikely, we are then faced with the dilemma that any lack of consistency between the two sets of responses may be at least partially attributed to actual changes in the structure of the group.

A further difficulty is encountered in connec-with restructuring groups in compliance with sociometric choices (Gottheil, 1952). If the group is changed, as orthodox sociometric practice demands, the retest is carried out on a new group and differences in the response data would, of course, be expected. If the group is not changed it might be argued that the ineffectiveness of the subjects' responses in producing significant changes in their interpersonal world alters their attitudes toward the sociometric questions so that one would again expect differences in response to the first and second administration of the technique.

We are thus partially opposing *test reliability* to *test sensitivity*. If we accept human groups as more or less constantly undergoing changes in their internal composition, only an insensitive test could give us a consistent measure from two administrations widely spaced in time. In such a circumstance changes in response may be evidence of the sensitivity of the test. Or, more convincingly, temporal consistency of the test, in these circumstances, would be evidence for the lack of sensitivity and representativeness of the data it provided.

Instead of repeat reliability, one may inquire into the *internal consistency* of the sociometric measure. Here one might in some arbitrary fashion divide the group in half, so that for any index each individual is given two scores, one from each half of the group. As P. N. Pepinsky (1949) has pointed out, reliability here becomes a measure of how consistently the individual is reacted to by the various members of the group, and no one has ever claimed that all members of a group ordinarily react to an individual in the same way. Thus, differences between the two halves are to be expected and are not evidence of the unreliability of the test. There are further difficulties involved in this approach because the average number of individual choices received in most groups is quite low. Dividing them in half leads to a large number of zeroes and small numbers, which in turn produces serious quantitative problems.

Finally, where multiple criteria are used in the test one may inquire into the consistency of choice patterns occurring in connection with the various criteria. This method corresponds roughly to the use of *alternate forms* in establishing the reliability of psychometric instruments. However, the consistent emphasis of sociometric investigators upon the crucial importance of the particular activity or criterion in terms of which the choice is made makes it difficult to equate the reliability of the test to the relation between different sociometric questions. Unreliability here may be a function not of the test's fallibility but rather an empirical demonstration that personal choice patterns vary directly with the activities in terms of which the choice is made.

The one circumstance when it may be appropriate to use the relation betwen separate sociometric questions as a measure of the internal consistency of the test is when the investigator is using a cumulative choice score which is made up of the combined total of choices received on two or more criteria. If in these circumstances it is assumed that the scores or choices on the various criteria are functionally equatable, it is perfectly legitimate to ask that there be consistency from one criterion to another.

To talk of the reliability of sociometric measures in general is like talking of the reliability of projective tests, questionnaires, or rating scales, without noticing that the range of reliability coefficients *within* any one of these techniques (depending upon the particular variety examined, the group being studied, the variables used, etc.) is much greater than the difference *between* these techniques. In reporting sociometric reliabilities it is a matter of great importance to specify the nature of the sociometric question used, the group being studied, the period between administrations (if a test-retest method is used), the particular score or index used, the number of choices required, etc.

In spite of the difficulties involved in studying the reliability of this instrument, there have been a large number of studies that provide information bearing on this issue. The interested reader should consult Bell and French, 1950; Bonney, 1943a,b; Byrd, 1951; Criswell, 1939, 1949; French, 1951; Gottheil, 1951; Hunt and Solomon, 1942; Jennings, 1950; Moreno, 1934; Newstetter, Feldstein, and Newcomb, 1938; Northway, 1952; and Zeleny, 1939. The results of these studies, and others, suggest the following tentative generalizations concerning the reliability of sociometric measures:

1. Most investigators report a relatively high degree of consistency in the sociometric pattern or sociogram over time, even though individual choices and rejections may fluctuate considerably (Moreno, 1934). It is important to appraise this generalization in terms of the data upon which it rests. As the analysis of sociograms is almost solely impressionistic (made by inspection) it is clear that this report is based upon impressions that investigators received when they examined two or more sociograms representing data collected from the same group at different times.

2. With somewhat better quantitative evidence, one might apply the same generalization to indices or scores derived from sociometric data (Criswell, 1939; Jennings, 1950). Although the particular choices and rejections may show considerable variation, the scores, e.g., choice score, emotional expansiveness, etc., show considerable consistency when observed over a period of time.

3. The reliability of the instrument seems somewhat greater when it is used with adults than when it is used with children. This is a rather ambiguous finding, as it is not clear whether the higher reliability in adults is mediated by greater stability in interpersonal relations or superior memory on the part of adults. Either interpretation seems consistent with other evidence.

4. There is some evidence that the least important or salient choices show the largest amount of change or unreliability. Thus, Moreno (1934) reports that following a three-month interval there was a shift in 8% of first choices and 18% of second choices, and this finding was in general replicated by Criswell (1939).

5. The stability of sociometric choices appears to increase with the passage of time during which the group has been in existence (Hunt and Solomon, 1942). There is a much higher incidence of change in sociometric choice during the early stages of group formation.

VALIDITY

Just as psychometric notions of reliability are difficult to apply directly to sociometric techniques, so too the conception of validity as "whether the test measures what it purports to measure" is not readily applicable. This difficulty in application is related to the question of whether the sociometric measure may be used in such a way that it purports to measure nothing more than verbal choice behavior, which it transparently does, and no further demonstration of which is needed. It is also related to the possibility that one may specify numerous "purports," all of which go beyond choice behavior and show considerable variation in the hands of different investigators.

Face validity. If the investigator limits his interest to a measure of interpersonal choice, it is clear that the somewhat dubious concept of "face validity" is applicable here. The sociometric measure is made up of interpersonal choices, and to talk of external measures or criteria to demonstrate that this is the case is to overlook the operational base of sociometric scores. Thus, where the investigator is willing to accept this limitation upon what he is measuring, it is clear that his problems are quite different from those of the test constructor who devises a set of items to measure some general disposition or underlying attribute. In the latter case there is a clear need for an independent demonstration of the test's sensitivity to the variable of interest. This difference between sociometric measures and many other psychological tests has been pointed to by both P. N. Pepinsky (1949) and Jennings (1950).

Data appropriateness. Even where the investigator is willing to limit his concern to sociometric measures as indicative of verbal choice behavior he does not altogether escape problems of evaluating the sensitivity of his measure. First of all, there is the question of systematic biases on the part of his subjects that may lead to deception in their responses. For reasons of dislike, resentment, or insecurity the subjects may be motivated to give an untruthful picture of the individuals they would actually like to interact with in particular activities. Second, although the subject may not be motivated to deceive he may be unmotivated to give careful responses. The sociometric answers may be insensitive either because the subjects did not fully understand the nature of the question or because they were not willing to devote the time and effort necessary to provide a thoughtful answer. This might result either from too long a questionnaire or from a general lack of interest in the investigation.

In both of these contexts the question of the sociometric questionnaire as a "motive" for the subject is worth stress. Moreno (1934) has emphasized the importance of this consideration and his demand that the sociometric test lead to restructuring the group is aimed at eliminating just such possible failures of cooperativeness or truthfulness.

Relation to independent measures. Sociometric measures are not used alone as indications of interpersonal choice patterns. In many cases investigators may equate an index derived from sociometric responses to morale, prejudice, social adjustment, or some other variable. In this case the relation of this measure to other measures of the same variable becomes a question of considerable interest. Failure to observe substantial relations is not, however, necessarily an indictment of the instrument's validity. If one defines minority group prejudice in terms of the extent to which an individual selectively chooses or avoids members of various minority groups it would not necessarily be an indication of the invalidity of this measure of prejudice to show that it did not correspond closely to prejudice as measured by one of the conventional paper-and-pencil attitude scales. Such a finding would represent an interesting question for formulation and further investigation, but it would not prevent any individual who so wished from continuing to define prejudice in terms of sociometric responses rather than attitude questionnaires. A similar consideration applies to differences between choice behavior as revealed in sociometric questionnaires and actual choice behavior as revealed in a study of overt behavior. There are important determinants or conditions which operate to determine overt behavior (physical constraints, social obligations, perceived likelihood of reciprocation, etc.) that do not operate so strongly in the anonymous verbal choice situation. Difference between behavior in the two situations is an interesting empirical finding but does not necessarily indicate that the one measure is invalid or insensitive. Rather it is confirmation of the fact that these two kinds of responses are differently determined, although closely related. Most of the reported studies dealing with this problem show moderately high correlations between sociometric choice and "actual" behavior, e.g., Byrd, 1951; Jennings, 1950; Danielsson, 1949. On the other hand, Polansky, Lippitt, and Redl (1950) report little relation between observed interaction and choices of whom one likes to be around.

The *multidimensionality* of the sociometric test makes the task of demonstrating its validity a rather unending one. Separate demonstrations would be necessitated or entailed for each variable which might conceivably be measured by the instrument.

Research utility versus validity. Actually, a far more reasonable question than that of whether a test actually measures the variable or variables it sets out to measure is the more general question of the extent to which it controls or relates to significant independent measures that are of interest to the theorist and investigator. Whatever it is eventually decided the test measures, if it relates successfully to a large number of pertinent, independent measures, it is clear that the findings will be of value and interest in developing knowledge in this area. In the following section we plan to deal at some length with this question and, so far as any meaningful estimate of the usefulness of the instrument is concerned, this is where one should look, rather than under studies specifically labeled validity studies.

In summary then, an examination of the validity of sociometric techniques is made complex by (1) the possibility of limiting one's measurement interest to verbal, interpersonal choice, in which case no validity demonstration is necessary; (2) the possibility of specifying a large number of variables which it is intended the test shall measure, in which case the job of relating sociometric response to independent measures becomes relatively unending; (3) the difficulty of interpreting failure of sociometric measures to show a high relationship to independent measures of the variables of interest.

RESEARCH UTILITY

The empirical value or usefulness of a measuring instrument can be appraised in many different ways. Here we propose to discuss the utility of our technique in terms of two major questions:

1. *What are some of the empirical concepts currently used by social scientists that can be defined in terms of sociometric operations?*
2. *What significant independent measures of behavior, or its environmental context, can be related to sociometric variables?*

Although these questions are designed especially to provide a representative picture of the actual and potential usefulness of sociometric instruments, similar questions might be asked concerning any psychological measuring instrument. If one views the main purpose of an empirical discipline as the building of stable, significant relationships that are explicitly interrelated (in the ideal case derivable from a general set of assumptions), then answers to the above questions can be expected to provide a relatively comprehensive picture of the likely contribution of any measuring tool.

EMPIRICAL CONCEPTS DEFINABLE IN SOCIOMETRIC OPERATIONS

Our purpose in this section is to discuss briefly some of the major variables of psychological interest that sociometric measures have been used to assess or appraise. Our consideration of the actual investigations that have employed sociometric measures to get at such variables will be highly schematic, as the number of relevant studies in many cases is extremely large. We shall be content here to give some impression of the general ways in which these measures have been used in connection with some of these variables.

Leadership. Although measures of leadership are legion in the psychological literature, virtually all of them have dealt with interaction between individuals, usually in a group setting. It is consequently understandable that many psychologists have found sociometric techniques useful in their studies of leadership. These instruments have assumed a particularly important role as psychologists have come to think of leadership as a function of the individual, the activity he is engaged in, and the group he is a member of, rather than as a quality that inheres solely in the person.

Perhaps the first sociometric study to address itself specifically to problems of leadership was Jenning's (1937) longitudinal study of a single cottage in a school for delinquent girls. Using as her sociometric criterion choice of eating compaion, she defined leadership as receiving five or more choices — this accounted for roughly twenty percent of the group. She discusses the constancies and the changes that appear in the sociometric pattern over a two and a half year period. Further reports by the same author (Jennings, 1947, 1950) describe in more detail the study of this same institution. As part of her extensive investigation she selected a group of overchosen and a group of underchosen subjects, the overchosen representing persons who were one standard deviation above the mean in the number of social choices they received, while the underchosen were one standard deviation below the mean. Approximately three months before the sociometric data were collected the girls in the institution had elected a House Council. From each cottage four council members were elected. In comparing sociometric choice with the election results Jennings used only the highest two, as these persons sometimes received such a preponderance of the votes that the individuals in the lower two positions were elected with very few votes. Examination of the two sets of data revealed that 90 percent, or eighteen of the twenty council members, were in the overchosen group. Further, the two persons who were not in this group fell only slightly below the necessary standard. The author concludes: "When allowance is made for the difference between being chosen from a community-wide base and being elected from the limited house population, it is evident that there is practically a one-to-one relationship between being elected to represent the house body in matters concerning the group and being chosen by community members on the sociometric criteria of living and/or working with them" (p. 38). Bell and French (1950) used sociometric measures to assess leadership in informal discussion groups. They found that changing the composition of the group and the discussion topic did not produce marked changes in the leadership standing of most of the subjects. Individuals tended to receive similar sociometric choice patterns even in groups with different members discussing different topics.

A number of investigations have been concerned with leadership in the military service as measured through sociometric devices. Wherry and Fryer (1949) examined the utility of sociometric ratings as criteria of leadership and concluded that they serve this purpose satisfactorily. Williams and Leavitt (1947) found that frequency of sociometric choice, made in terms of the criteria of "all-round ability as a combat officer," was positively associated with ratings of combat performance. Stogdill (1949), in a study of Naval officers, defined leadership in terms of the number of times an individual was mentioned by other officers as a person with whom they spent time "getting things done." When this measure was related to other measures it was found to correlate with military rank, methods used with subordinates, and other measures concerned with how the individual used his time. The sociometric data, when plotted graphically, also produced interesting information concerning the informal relations existing within the formal structure of the group. Gibb (1950) proposes a comprehensive definition of leadership and then sets out to examine the relation between ratings of this dimension and three kinds of sociometric response. The three sociometric questions employed the criteria of spending leisure time with, working with, and the person whose removal from the group would bring about the largest group change. Two sets of ten-man leaderless groups (college students and Army officer candidates) were studied. Observer ratings correlated in the neighborhood of .40 with sociometric choice for the criteria of spending leisure time and working together. Sociometric choice for the third criterion, however, correlated about .80 with the leadership ratings. Gibb's findings indicate both the imperfect association between sociometric choice and some aspects of leadership, and also the importance of the particular criterion that is used as the basis for the sociometric choice.

Usually when the sociometric investigator is interested in leadership he has employed criteria that direct the respondents to think of leader functions, e.g., "Who would you choose to lead your squad on a combat mission?" In some cases, however, the investigator has attempted to infer information about leadership from sociometric questions not directly related to leadership, e.g., choice for roommate, eating companion. As Gibb's findings imply, not all sociometric criteria are equally appropriate for the measurement of leadership, and the person who uses remote criteria may secure a relatively poor measure of leadership.

A further difficulty involved in sociometric measures of leadership is the possibility that the actual leader of the group may not be distinguished by the frequency with which he is chosen by other group members. Moreno (1934) described the individual who secured a few key choices and consequently occupied an influential position in spite of the fact that his choice status was quite low. Gibb (1950) has pointed to the distinction between the individual who is the social center of the group and the person who is the functional leader of the group. In some cases the leader and the person who has the highest choice status are the same, but this is by no means always the case. Hunt and Solomon (1943), in a study of summer campers, report that individuals who stand out as leaders in the judgment of camp counselors are high in social choice but not clearly distinguishable from the rest of the group on these grounds. Katz's (1953) recent proposal of a status index that takes into consideration both the number of choices and the identity of the choosers may prove the most efficient means of meeting this difficulty. The reader will find a more detailed discussion of leadership research, including sociometric studies, in Chapter 24.

Our discussion implies that there are many pitfalls involved in the sociometric measurement of leadership. Further, it seems clear that any comprehensive study of leadership necessitates more information than can be supplied by this technique alone. However, sociometric measures can provide important data bearing on leadership and in certain circumstances can be used directly to measure this variable.

Minority group prejudice. From the very outset of sociometric measurement there has been interest in applying this technique to the study of group cleavages and prejudice. Although in some instances sociometric measures have been used to assess individual prejudice, they have been used much more frequently to examine groups in search for cleavages that are established along minority group lines.

In *Who Shall Survive?*, Moreno (1934) reports a study of public school children in which minority group cleavages were observed only beyond the fourth grade. He also introduces the notion of a "saturation point," a point beyond which introduction of further minority group members will produce cleavages and other group changes. Criswell (1942) has discussed this concept and concludes that it is imprecise and difficult to measure, and so should be used only with great caution.

The first study to focus primarily upon minority group cleavages was Criswell's (1937, 1939) investigation of Negro-white relations in a large number of urban classrooms ranging from kindergarten to the 8th grade. She measured cleavage by comparing the choices directed within the "own" ethnic group to the frequency of choices that would be expected by chance. Initially, she found that sex appeared to be a much more important determinant of cleavage than ethnic group membership, and consequently she limited her group comparisons to individuals of the same sex. Her findings, in general, support Moreno's, as there appeared to be little cleavage between the Negro and white children prior to grade three, and by grade five the cleavage had reached its height. She also reported that withdrawal of one group is not always associated with withdrawal by the other group, and further that the nature of the cleavages observed was influenced by the relative proportion of white and Negro students present in the class. A detailed discussion of methods of measuring group preferences using sociometric data is supplied by Criswell (1943). A study by Loomis (1943) employed a modified sociometric technique to study cleavages in two southwestern high schools. He found that Spanish-American and American students reported much greater association with members of their own groups. The minority group, whether American or Spanish-American, showed a stronger tendency toward in-group selection and exclusion of the out-group than was shown by the majority group.

Northway (1946) has indicated her belief that the sociometric technique can be used effectively to depict prejudice, and her sector target diagram (Northway and Quarrington, 1946) provides one means of examining minority group relations through sociometric data. The preliminary findings she reports suggest that social status varies with the ethnic group of the individual and that the size of the minority group has an important influence upon group relations. A novel approach to ethnic group cleavages, using the relational analysis, is reported by Goodnow and Tagiuri (1952). They found that religious groups tended toward in-group choice and, further, that social status varied with the group membership. Most interesting was the finding that when asked to guess who would choose them, the subjects showed a high degree of accuracy in their predictions. Thus the rejected minority group appeared to be well aware of its status in the group.

The use of sociometric response to provide an individual measure of prejudice is described by Zeleny (1947a). He asked members of a class to make choices among the class members for membership on a committee established to combat prejudice. The "prejudice index" was derived from the number of times an individual was chosen or rejected by the other members of the class — the higher the index, presumably, the lower the prejudice. In the same study, Zeleny related his prejudice index to scores provided by a questionnaire designed to measure willingness to accept Negroes and Spanish-Americans, and he found a moderate positive correlation between the two measures. Zeleny's method is somewhat cumbersome and indirect, since he depends upon the judgments of peers as to how prejudiced the person is. Further, the choices for committee members are bound to be influenced by personal considerations that have no relation to prejudice.

Where the sociometric technique is administered to groups containing appreciable numbers of one or more minority groups, a much more direct and satisfactory measure of prejudice can be obtained. Various ratios or indices representing the relative choice by each individual of minority and majority group members can be obtained. Such a measure possesses obvious advantages because it gets away from the abstract, class notion of the minority group member. Instead of asking the respondent to verbalize his attitudes toward the Protestant, Jew, Catholic, or Negro, he is asked to select those individuals with whom he would like to associate in various activities, and perhaps those with whom he would not like to associate. A measure based on this kind of evidence represents not the subject's generalizations concerning his relations with minority group members, when they have been identified as minority group members, but rather an indication of whether this person actually uses minority labels as criteria in the process of determining with whom he will interact.

Thus, sociometric measures may provide a convenient means of estimating the extent to which *groups* are divided or disrupted by withdrawal from minority members. They also permit, in certain circumstances, a sensitive portrayal of the extent to which prejudice is determining an *individual's* personal choices and rejections.

Social adjustment. One of the measurement possibilities inherent in sociometric response is estimation of the effectiveness with which an individual interacts as a member of a group. The success of his interaction (social adjustment) may be equated to various choice and rejection measures obtainable through sociometric techniques. There are obvious difficulties involved in such a procedure, one of these being the danger of equating this measure to personal adjustment or freedom from conflict, anxiety, or disrupting psychological problems. Although an individual's social adjustment may be expected to bear considerable relation to his personal adjustment, the two variables are not the same. In particular cases there may be striking discrepancies between them, e.g., the basis for an individual's social success may be a lack of personal security that makes it impossible for him to countenance rejection or loss of love.

The use of sociometric methods to measure social adjustment is implicit in Moreno's original formulations. While he did not refer directly to his measures as indicators of social adjustment, this area of interest is involved in much of his writing. He implies that low choice status or high rejection status is evidence that the adjustment of the subject is not good. Concepts such as "isolate," "neglectee," and "rejectee" clearly connote adjustment.

Jennings (1943, 1950) points out that description of the individual as an "isolate" refers to the choice behavior of the group and provides no evidence of whether the isolated individual is isolated of his own accord or involuntarily. Similarly, being overchosen does not necessarily imply awareness of this on the part of the overchosen, or that he is "happy" because of this status. Jennings suggests, reasonably, that the desirable sociometric pattern should be decided through consideration of the particular setting in which the choices are being made. She also implies that the mutuality and stability of choice are important considerations in any question of this sort. Northway (1944) has suggested distinguishing between social "acceptability" and "acceptance." She suggests that sociometric data provide evidence of the "acceptance" of the person within a specific social setting. The social "acceptability" of the person can be measured only with a much wider range of information concerning reactions of people to this person in a variety of social settings. While this is a legitimate caution, it is important to realize that a single set of sociometric responses is no better or worse than a single set of observations of any other kind. The problem of generalizing findings is always present and is no different with sociometric data than elsewhere.

A number of studies have shown that aspects of sociometric response are related to measures used to estimate personal adjustment. Many of these studies will be summarized in a subsequent section of this chapter.

Our discussion has implied that the investigator who wishes to do so may use sociometric results directly as a measure of how well the individual adjusts to or is accepted in a social setting. However, it is dangerous to assume that such a measure reflects perfectly the personal adjustment of the individual, although there is considerable evidence to suggest *some* relation between these two variables.

Social sensitivity. In assessing the quality of interpersonal relations an important consideration is the accuracy with which an individual can identify the reaction he arouses in others. Presumably, the well-adapted person should be able not only to arouse positive reactions from others but should also be aware of this state of affairs, and ideally he should be able to identify those persons from whom the choices stem.

The use of the sociometric self-rating or "relational analysis" provides a ready means of assessing this variable. The investigator may simply ask the respondent to indicate the general reaction of the remaining members of the group or he may ask for estimates of the individual reactions of all other members of the group. Tagiuri (1951) has shown that the accuracy with which an individual can identify the reactions he produces in others is related to personal adjustment as well as to a complex of factors having to do with minority group membership and the particular social setting in which the judgment is made.

Social status. The concept *social status* has been associated with sociometric measures in at least two contexts. First, choice status under several definitions has been directly equated to social status in the group, and second, some investigators have attempted to demonstrate that high choice status would be associated with a favorable position in the general status structure of the community.

Of the many investigators who have used sociometric response to measure social status, Zeleny (1940a, 1941) was perhaps the first to propose a specific sociometric index under this label. He proposed that social status be measured by the sum of choices times an intensity score divided by $N - 1$. More recently, Lemann and Solomon (1952) recommended measuring social status through the use of choice and rejection. Such indices have the advantage of not relying solely upon positive choice but rather including rejections and (reciprocally) indifference or no response.

There is excellent evidence that social status as measured by sociometric measures is related to other indices of the individual's group status. A number of studies to be discussed later in the chapter (Bonney, 1944a; Loomis and Proctor, 1950; Lundberg and Steele, 1938; etc.) agree in indicating that socioeconomic status is positively related to sociometric status. Further investigations (Lundberg and Beazley, 1946; Bonney, 1946) suggest that socioeconomic cleavages are accompanied by sociometric cleavages, or at least that there is a tendency for individuals choosing one another to be of similar socioeconomic status. Other measures of social status in the community, such as occupation (Loomis, 1946) and observer's judgments (Powell, 1951), also show association with sociometric data.

Loomis, Beegle, and Longmore (1947) have presented a detailed critique of the concept of social class or social stratification and propose a procedure for testing empirically for the existence of social classes. They suggest using a modified sociometric technique (family visiting patterns) to provide a means of finding out whether, within the particular group studied, individuals tend to cluster or form interpersonal cleavages corresponding to the classes that have been used in describing the group. A study by Longmore (1948) illustrates the application of this method in a concrete setting.

The use of sociometric data to provide information concerning individual social status and social class cleavages appears both feasible and defensible. It is, of course, important to realize that there are many other sources of data that are important for a full representation of social status.

Morale. Morale is considered an important dimension in the description of groups, and aspects of this dimension may be approached through sociometric techniques. Zeleny (1939) was perhaps the first investigator to attempt a direct measurement of morale through the use of sociometric data. He defined morale as the "shared feeling of like among members of a group" and proposed that it could be measured by means of the *morale quotient,* which was the number of choices a person received, times an intensity score divided by the total possible number of choices. This index could be computed either for the group as a whole or for the individual member. Zeleny was able to show that this index possessed reasonable stability on successive daily administrations and that it also correlated to a moderate and significant degree with the results of a questionnaire measure of morale. He demonstrated that, given the sociometric data, it was possible to reform classroom groups so as to produce sizeable increases in the morale quotient.

Several of the studies to be discussed later in connection with military applications (Jenkins, 1947; Goodacre, 1951) have found sociometric measures useful indicators of morale. Most of these studies have defined morale in terms of the relative amount of in-group choice, with some emphasis upon the existence of positive choice of the formal leaders by members of the group.

Again we find that an empirical concept widely used in social psychology may be defined either partially or wholly in terms of sociometric data. It seems likely that the information provided by sociometric measures concerning group morale is at least as valuable as that procurable through the use of customary questionnaire devices.

Group structure. The emphasis of most sociometric research has been upon studying group structure, and a large number of measures have been used in describing this phenomenon. The various methods of matrix analysis, index analysis, and sociographic analysis are all assumed to be alternate means of intercepting or representing this structure. Concepts such as clique, subgroup, isolate, chains of communication, index of cohesiveness, etc., are employed as means of describing specific aspects of the structure.

Moreno (1934) dealt with group structure by comparing the relative distribution of scores or indices in different groups. For example, he examined the percentage of isolates, mutual pairs, triangles, and chains in different grades of schools and found that the grades could be distinguished in terms of these scores. He also showed that these same grades could be differentiated in terms of sex and minority group cleavages. The description of groups in terms of cleavages has been an object of considerable interest. The work of Criswell (1939; 1943) in measuring cleavages, group integration, and group preference, in addition to various publications of Loomis (1943, 1946), is among the best known investigations in this area. Bronfenbrenner (1944a) derived an index of group coherence from the number of mutual choices and found that the degree of coherence increased from kindergarten to the upper grades. He also found that the amount of differentiation within the group (cliques and cleavages) increased with the higher grades. Many other studies have used these and other measures derived from sociometric data to describe and classify groups, and some of these are considered in Chapter 22.

Although sociometric data can be used to infer important information concerning group structure, it is obvious that many other sources of data are important for the full description of these structures. Information derived from direct observation of the group, the formal organization of the group, the occupational and avocational associations within the group are all of considerable importance in this context.

SIGNIFICANT MEASURES RELATED TO SOCIOMETRIC VARIABLES

Information concerning the range of variables to which sociometric response has been related is of interest both as a means of indicating areas of investigation in which the instrument may be useful and as an index of the general fruitfulness of the instrument. If a measure can be shown to relate to a number of significant variables, no matter how vacuous or psychologically uninteresting the content of the measure may appear, it will be of interest to many investigators.

We have chosen to divide studies that have related sociometric responses to independent measures in terms of whether the study emphasizes demographic, cultural, intellectual, performance, attitude, or personality variables.

Demographic variables. There are numerous studies that show some aspect of sociometric response to vary with age, sex, religion, socioeconomic status, or educational level. We limit our interest here to a representative group of such studies.

Early in the use of sociometric measures it was found that *socioeconomic status* in many settings showed a positive correlation with sociometric choice and, further, that there was some tendency for individuals to choose other individuals within the same socioeconomic class. A study of personal relationships in a small New England village by Lundberg (1937), Lundberg and Lawsing (1937), and Lundberg and Steele (1938) revealed that socioeconomic status was positively related to the frequency of sociometric choice and that individuals tended to choose individuals who were slightly higher in socioeconomic status than the chooser. Another community study, conducted by F. A. Stewart (1947a) in a southern town, indicated that individuals who were chosen as most influential tended also to be the wealthiest. Investigation of a group of fifth grade students by Bonney (1944a) indicated that sociometric status was positively related to socioeconomic status. Grossmann and Wrighter (1948) found that in a sixth grade classroom socioeconomic status as measured by the Barr Scale was related to social status. An investigation of agricultural county agents and librarians who were participating in a university workshop indicated that sociometric status was positively related to annual income (Loomis and Proctor, 1950).

Evidence for the tendency of individuals to choose other persons of the same socioeconomic background was supplied by Bonney (1946a), who reported that among fifth grade students mutual sociometric choice is accompanied by similarity in socioeconomic status. In a similar study conducted in a woman's college, Lundberg and Beazley (1948) used two measures of socioeconomic status (tuition reduction and self-support) and found that, using either measure, individuals tended to choose other persons within the same socioeconomic bracket. Using family visits to permit sociometric representation of a German village, Loomis (1946) was able to show that occupational role was an important determinant of sociometric cleavages.

Related to socioeconomic status is the question of association between *family size* and sociometric variables. A series of studies carried out by Bonney (1942, 1943b, 1944, 1949) have indicated that the number of siblings appears to influence social choice, particularly at the extremes. Thus individuals who are "only" children appear to secure more than their share of positive choices, while at the college level individuals who have six or more siblings tend to receive less than their share of positive choices. Loomis, Baker, and Proctor (1949) found little evidence of any relation between family size and social status in a group of college freshmen.

There is relatively clear evidence that *religion* and *ethnic group membership* are important determinants of choice status as well as clique formation. Moreno (1934) and Criswell (1939) both report convincing evidence that sociometric cleavages are established with minority group membership as a primary determinant. A recent study by Goodnow and Tagiuri (1952) shows both the relation between social status and religious-ethnic grouping and also the tendency to form cleavages around these distinctions. Lundberg and Steele (1938) found membership in the same church to be one of the most important determinants of clique structure in a small New England village. Bonney (1949) found no relation between church membership and the number of choices received in a large college sample. He did find, however, that those individuals who were not affiliated with any church received significantly more choices than they would be expected to by chance alone. Further, six of the seven groups Bonney studied showed significant tendencies toward in-group selection — they made more choices within their own religious group than would be expected by chance.

Beginning with Moreno's early investigations, there has been an increasing amount of evidence that *sex differences* are related to both social status and clique formation. One of the earliest and most enduring of the group cleavages that appear in our society is the one based on sex (Moreno, 1934; Criswell, 1939). A group of teen-age campers are reported by Faunce and Beegle (1948) as showing marked sex cleavages. Further, there is interesting evidence that social status, depending upon age and the particular group, is related to the sex of the individual. F. A. Stewart (1947b) reports that in a southern town women selected slightly more men than women as influential, while men selected almost 30 percent more men than women. This difference was observed in spite of the fact that subjects were asked to name both men and women. Several studies by Bonney (1942a,b, 1944b) have suggested differences between sexes in elementary school children, generally showing girls to have higher choice status.

A number of studies, e.g., Lundberg and Beazley (1948), Lundberg, Hertzler and Dickson (1949), Festinger, Schachter, and Back (1950), have shown that choice varies with *physical proximity*. Individuals who are near to one another are more likely to choose each other than persons who are more remote. Danielsson (1949) reports that in a nonliterate society rejection is positively related to physical distance, e.g., the most distant groups within the prescribed area had the highest percentage of rejection. However, in none of these studies was acquaintance or degree of contact controlled.

Level of *education* frequently appears as a sociometric correlate. F. A. Stewart (1947a) concludes that in Southtown education was "virtually a prerequisite" for occupying an influential position. In a study of college students, Bonney (1949) reports that those individuals who had been in school for less than one semester received significantly fewer than the expected number of choices, while students who had been in school for more than six semesters obtained significantly more than the expected number of choices. Vreeland (1942) had fraternity members rank other members of their group in terms of how friendly they felt toward them. Results indicated a strong tendency for individuals to direct their preference within their own class and, further, individuals in upper classes were given higher ranks than those in lower classes.

Age also is known to show association with sociometric variables. F. A. Stewart (1947a) found in his community study that individuals selected as influential tended to be older than individuals not so selected. Moreno (1934) reported characteristic differences in sociometric patterns that could be expected to appear with increasing age from shortly after birth on. Both Moreno (1934) and Criswell (1939) have provided evidence that various kinds of group cleavages are especially likely to occur in particular age groups. Faunce and Beegle (1948), studying adolescent summer campers, found that age was a significant determiner of cleavages.

Cultural variables. The application of sociometric techniques to cross-cultural study is still in an embryonic stage. The possible value of such a union has been considered by several authors (Mead, 1947; Ashley-Montagu, 1945), but only a small number of anthropological investigations have been reported that make use of sociometric procedures. Perhaps the best example of these is Danielsson's (1949) study of attraction and repulsion among the Jibaro Indians, a South American head-hunting tribe. The author examined three subgroups within the tribe and also compared some of his findings with those reported by Lundberg (1937) from observation of a Vermont village. Danielsson found that for sociometric variables such as cohesiveness and interaction the three subgroups differed markedly among themselves and also differed from Lundberg's data from our own society — the Jibaro groups showed a higher level of interaction. The author diagrams the sociometric position of various "key" figures in the subgroups and discusses these results in view of the information he has about the status and function of these individuals. In brief, the investigator finds that sociometric procedures serve a useful function in a fieldwork setting. Loomis and Powell (1949) used sociometric measures in comparing class structure and friendship patterns in two Costa Rican communities. One of the communities was characterized by peasant-owned farms, while in the other community most members of the community were employed on a large estate. The two groups differed not only in class structure but also in terms of such variables as the relative number of sociometric choices directed outside the group and the number of cliques formed within the group. A sociometric comparison of stable and changing neighborhoods in two large cities by Deutschberger (1946) revealed specifiable subcultural differences. Individuals in changing neighborhoods made significantly fewer choices and directed these choices more to their own neighborhood.

Intelligence and performance. Most studies that have related sociometric variables to intelligence have found some relation between status and intelligence. Thus, Bonney (1944) found that sociometric choice in a group of elementary school children was positively related to intelligence as measured by group intelligence tests, and he concludes that the upper quartile in social choice was distinctly superior intellectually to the lower quartile. This same finding held for third, fourth, and fifth grades. An investigation of a group of sixth grade students carried out by Grossmann and Wrighter (1948) related intelligence as measured by the Stanford-Binet to choice and rejection. The authors report that the group that was high in sociometric status was significantly higher in intelligence. They suggest, however, that high intelligence does not necessarily increase the likelihood of choice over that of the person of average intelligence.

There is slight evidence that individuals may choose other individuals who are similar to them in intelligence. Bonney (1946) found that a group of elementary school children who mutually chose each other were more similar in intelligence, as measured by group tests, than a corresponding group of subjects who did not choose each other. However, Lundberg and Beazley (1948) found little tendency for female college students to select other students who were similar in ability as measured by an academic aptitude test.

If achievement and *performance* are divorced from economic factors, there is only limited evidence indicating association between sociometric variables and these factors. Bonney (1946) found very little relation betwen friends in their academic achievement as measured by a reading test. Lundberg and Beazley (1948) found no significant tendency for female college students to select individuals in their own scholarship group. On the other hand, a study of sixth grade students conducted by Grossmann and Wrighter (1948) revealed sociometric status to be related to reading ability, at least to the extent that individuals in the lowest quarter of their group in social status were significantly lower in measured reading ability than those in the upper quartile of social choice. Interested in a different kind of performance, Hunt and Solomon (1942) found that counselor ratings of athletic ability correlated .60 with a sociometric measure of status in the first week of a summer camp but dropped to only .25 by the eighth week of the camp. Smith (1944) found a slight tendency for individuals engaging in similar athletic activities to select each other more often than would be expected by chance. Again most of these studies do not control for degree of contact or acquaintance.

Attitude and value. In view of the preoccupation of many social psychologists with the measurement of attitudes it is no surprise to find that many studies have been conducted exploring relations between sociometric variables and attitude measures. In general, this research suggests that within certain groups maintenance of particular attitudes is correlated with sociometric status and, further, that individuals who choose one another tend to be similar in their attitude structure.

One of the most interesting of the studies concerned with attitudes is Newcomb's (1943) attempt to assess the role of personal and community factors as determinants of attitudinal shifts in a small college for women. He found that liberal political and economic views were positively associated with sociometric choice. This was equally true when the choice was made in terms of "community prestige" and when it was made in terms of an individual to serve as an enduring friend. When individuals were assigned "community reputation" scores by means of a "guess who" technique, those individuals who received the highest scores were the least conservative. An unpublished study by Smelser (1952) related attitude measures to sociometric dimensions in a sample of six female college dormitories. His attitude questionnaire measured extremeness or conservatism of attitudes in eight areas: dominance-submissiveness in the female courtship role, career versus marriage, sexual freedom versus sexual restraint, motherhood, glamour, political-economic conservatism, religious conservatism, and ethnocentrism. When the results for the six dormitories were combined there was no evidence of any association between choice status and the eight attitude dimensions measured. When the dormitories were considered individually, however, there was evidence for a significant relation between all of these attitudes and social status, although the direction and magnitude of the correlations varied considerably between the different dormitories. Smelser concluded that there were separate norms or social climates that operated within the dormitories and that these were much more important in determining social status than any college-wide norm. He also divided his groups into cliques, using the method of Forsyth and Katz (1946), and found his cliques to be generally homogeneous in their attitude structure.

One of the earliest studies concerned with the similarity of attitudes among friends was Winslow's (1937). He administered a long questionnaire to a group of 86 students and asked each of them to have a friend fill out the questionnaire. When the scores of friends were correlated for measures of attitudes toward the Negro, American foreign policy, current economic policy, religion, and domestic political policy they were found to be significantly related in a positive direction for all but attitudes toward foreign policy. Even here the relation was positive and approached significance. A study by Richardson (1940) used observational data to determine patterns of friendship and found that friends, both female college students and adult women, tended to resemble one another in values as measured by the Allport-Vernon Study of Values. Bonney (1946) reports a significant positive relation between the interests, as measured by the Kuder Preference Record, of individuals choosing each other. Using an interest inventory for elementary school children, he found little evidence for any similarity in the attitudes of individuals choosing each other. If we accept the subject that an individual selects for his college major as implying certain interests or attitudes, there are several studies that demonstrate definite in-group choice tendencies on the part of individuals who have made the same educational choice, e.g., Lundberg and Beazley (1948), Lundberg, Hertzler, and Dickson (1949). It is difficult to interpret such findings, as it is clear that other factors, such as physical propinquity, operate in addition to the possible attitudinal similarity. In a postwar study of a German village, Loomis (1946) has provided evidence that sociometric cleavages may be at least partly determined by political attitudes, as inferred from membership in the Nazi party.

Personality. It is a matter of absorbing interest to both professional and lay audiences that particular patterns of interpersonal relations may be accompanied regularly by specifiable motives or aspects of personality. The study of motivational variables in relation to sociometric measures has been characterized by many of the shortcomings that have been manifest in other areas where motivational analysis has been a topic of recent interest. A major weakness derives from a reliance upon measures that are easy and inexpensive to administer but whose sensitivity to motivational variables is highly suspect. A further deficiency has resulted from the tendency to explore relations between sociometric findings and personality attributes "at random," with little or no attempt to engage in conceptual analysis in advance of investigation to permit prediction of the kinds of variables that may be expected to associate with sociometric data.

Among psychologists, measurement of personality variables through the use of *rating scales* is an old but not especially honored method. The most important questions to be asked of any study using this approach are the care with which the variable has been defined, the adequacy with which the raters have been trained, and the certainty that the raters have been kept from knowing of the variables to which their ratings are to be related. Very few of the studies to be discussed have answered these questions satisfactorily.

In a recent study of three small women's dormitories in an urban college, Lemann and Solomon (1952) studied six personality variables. All of the variables were bipolar and the investigators divided them into two categories: alpha, those ratings that ran from good to bad (generous-stingy, affectionate-cold, enthusiastic-apathetic); and beta, those that ran from bad to good to bad (dominating-submissive, shy-bold, stubborn-yielding). Their findings indicated that the three status groups were clearly differentiated on the alpha scales but not on the beta scales. The high in social status, as measured by a combined choice and rejection score, were characterized as generous, enthusiastic, and affectionate. When the subjects were grouped in terms of their noticeability or their indifference score (the number of combined choices and rejections subtracted from the total number of responses possible), a relationship was found between this division and ratings received on the beta scales. Individuals low in noticeability are rated as dominating, bold, and stubborn. The investigators also secured a measure of insight based on the discrepancy between self-ratings and other-ratings for the same variables. This index did not appear to be related to social status.

It is apparent that having the same individuals perform the personality ratings and the sociometric ratings poses certain problems in the interpretation of any association between these variables. Lemann and Solomon suggest: "It may be that a general 'halo effect' surrounding girls with high status causes them to be rated favorably on most traits" (p. 35).

In a similar investigation, French and Mensh (1948) studied 34 college students who were members of the same sorority. The subjects were asked to make three choices and three rejections for roommates and also rated each other on punctuality, sociability, fair-mindedness, intelligence, self-confidence, and sense of humor. Social status was measured by dividing the group by inspection into the majority ingroup and nine isolates. The majority group was then further subdivided into those who received rejections and those who did not. This resulted in three groups of varying social status. A comparison of these three groups in the personality ratings they received indicated that, in general, the individuals high in choice status were rated highest on those personality variables that are normatively most valued.

In considering the association they found between social status and rated personality variables the authors suggest a number of possible explanations: "Do individuals attain high status because they manifest these traits? Or do they develop these traits in consequence of having attained high status? Or are they merely thought by raters to stand high in these respects because they have high status or are admired as individuals? Or does high status permit the conspicuous display of these traits without there necessarily being any fundamental personality differences?" (p. 344)

The findings of these two studies are both interesting and consistent with what we would expect if personality factors were closely related to social choice. However, the interaction between sociometric ratings and personality ratings makes such findings of little significance as evidence of the relation between interpersonal relations and personality variables. To secure data that will be pertinent it is necessary that the personality ratings be provided by someone other than the individuals making the social choice.

Beginning with Seagoe's (1933) investigation, there have been a number of studies which have shown that sociometric choice is related to behavior ratings secured through "guess who" techniques. Although these studies measure personality somewhat more indirectly than ordinary rating methods, they nevertheless seem open to the same criticisms that have been leveled at the studies discussed above.

Probably the majority of studies concerned with personality and sociometric choice have relied upon paper-and-pencil *questionnaires* for measurement of personality dimensions. The following investigations are quite typical of a much larger number that have been carried out during the past twenty years. In a study conducted prior to the public description of the sociometric technique, Flemming (1935) required a group of students to nominate their best friends and in addition fill out questionnaires designed to measure such variables as pleasingness of personality, steadiness of emotional response, social adjustment, introversion-extraversion, and social intelligence. The reported findings indicated that friends were similar to each other in most of these characteristics, especially pleasingness of personality and social adjustment. An early investigation of a large number of elementary school children by Pintner, Forlano, and Freedman (1937) used questionnaires to measure ascendance-submission, extraversion-introversion, superiority-inferiority, and emotional stability. In addition, the children were asked to name their three best friends. The authors found no relation between friends in the measured personality characteristics. Van Dyne (1940) investigated the relation between personality factors and friendship in a group of 42 female summer camp members. The campers were asked to indicate the five persons they would most like to have as friends, and they also filled out the Bernreuter Personality Inventory. Correlational findings indicated that girls tended to choose as friends other girls who were similar to them in dominance and stability. Although the correlations for these two variables were significant they were very low in absolute magnitude. There was no evidence of any similarity between friends for the variables of self-sufficiency, emotional stability, introversion, and self-confidence. A group of 80 children in an elementary school were studied by Bonney (1944), using the California Test of Personality and a multiple criteria sociometric questionnaire. He found the total adjustment score of the inventory correlated .49 with sociometric status. Consistently, when the subjects were divided into quartiles in terms of social choice, the quartiles were arranged in a regular order with the high in social choice also the highest in mean adjustment scores. A more recent study by Bonney (1946) showed little association between friends in their scores on the California Test of Personality in either elementary school or college groups. There was, however, a moderate positive correlation between friends in their scores on the Social and Emotional dimensions of the Bell Adjustment Inventory. In a similar study, Grossmann and Wrighter (1948) found that sixth grade students who were very high in sociometric status secured much higher total adjustment scores on the California Test of Personality than did a similar group of students who were very low in sociometric status.

The upper and lower quartiles in social status of a large number (692) of ninth grade students were compared by Kuhlen and Bretsch (1947) in their responses to the Mooney Problem Check List. The findings indicated that the low in social status tended to check a significantly larger number of problems as "often" present, but the two groups were not differentiated in the frequency with which they checked problems as "sometimes" present. The pressing problems of the low in social status centered around social skills, lack of status, family problems, and dislike of school. A large number of fifth and sixth grade students were divided into three social status groups by Baron (1951) and compared in their responses to a mental health inventory. In general, the inventory appeared to differentiate between the three status groups, with the low in social status making more "unfavorable" responses than either of the other groups. Lindzey and Urdan (1954) related measures of security, dominance, happiness, and achievement derived from self-ratings, questionnaires, and a sentence completion test to various sociometric measures, including sociometric status and clique structure. They concluded that the questionnaire measures showed much stronger relations to the sociometric indices than either of the other two types of measures, although there was some reason for questioning the adequacy of the projective measure of these variables. They also found evidence for a curvilinear relation between social status and the variables of dominance and security, i.e., high social status was associated with median security and dominance.

Although the findings of these studies are not altogether consistent, there does appear to be some evidence that personality factors as meassured through self-inventories relate to aspects of sociometric response.

A small number of studies have attempted to measure personality factors through *interviews*, collection of *observational data*, or some combination of the two methods. Perhaps the first of these was reported by Richmond (1936). In her paper she indicated simply that when student nurses were interviewed after sociometric data had been collected, the impressions they created in the interviewer were consistent with the interpersonal data. Another study in this area, carried out by Northway (1944), was focused upon children who were very low in choice status. She collected personality data through interviews and observation of children in the fifth and sixth grades who were in the lowest quartile in social choice. This data suggested to Northway that there were really three distinct types of "outsiders," or individuals low in social status. The first group was made up of the "recessive children" who were low in energy, possessed little or no drive, and in general created a picture of listlessness and general lack of motivation. Second, there were the "socially disinterested children" who were less defective in general energy and motivation but who did not seem interested in gaining social approval. Finally, there were the "socially ineffective children" who were characterized by aggression, boastfulness, and other asocial responses but who appeared strongly motivated to secure social acceptance. In a study of over 400 adolescent delinquent girls Jennings (1947) compared the overchosen (1 standard deviation above the mean in frequency of choices received) and the underchosen (1 standard deviation below the mean). She concluded from observational and interview evidence that the overchosen showed a large number of distinctive qualities, among which were ability to establish rapport quickly and effectively, fairness, exhibiting anger primarily towards those they consider "should know better," securing more and more responsibilities for members of their groups. The underchosen more frequently interfered with the group activities, showed little initiative in starting projects or promoting understanding among group members, and in general seemed to impede rather than facilitate interpersonal relations.

If we consider accident proneness a personality variable, there are several studies that have used observational evidence to relate this variable to sociometric indices. A study by Fuller and Baune (1951) of third grade students found that when the subjects were divided into upper, middle, and lower thirds in terms of social choice there appeared to be a descending order of accident proneness; i.e., the most popular children tended to have the smallest number of accidents. Not only did the children high in sociometric choice tend to have fewer accidents in general, but those accidents they did have were primarily nonsocial, that is, they did not occur while interacting or playing with other children. On the other hand, the children low in social choice tended to have a relatively high incidence of social accidents. A related study was carried out by Speroff and Kerr (1952) with a group of 90 steel mill workers. They found that a social status score based on sociometric choice and rejection showed a highly significant relation with the frequency of industrial accidents, the low in social status having accidents more frequently.

Several relatively well-controlled studies have attempted to relate sociometric response to personality factors as measured through the use of *differentiated groups.* A recent study by French (1951) attempted to measure adjustment through psychiatric referrals, sick-bay attendance, and disciplinary offenses, and related these indices to a sociometric measure of social status. In the large group of Navy enlisted men studied, sociometric choice was negatively related to sick-bay attendance and disciplinary offenses. The number of psychiatric cases was not large enough to provide an adequate test of the relation to sociometric data, although the evidence observed was in the expected direction, suggesting that those individuals who had undergone psychiatric examination were relatively lower in social status. Tagiuri (1951), employing the method of relational analysis, compared three groups of prep school students — those who were successful (good athletes, high grades, officeholders), average students, and maladjusted students (seeing a psychiatrist). He found that the maladjusted students were characterized by low social choice and a small number of guesses that others would choose them. Further, the successful students tended to be higher than the average students on both of these variables.

The use of *projective techniques* for measuring personality variables to be related to sociometric measures is just beginning to attract attention. The relative paucity of studies involving this method of measurement, widely considered most sensitive to covert aspects of motivation, is in large part attributable to the relative effort involved in the collection and analysis of such data and the small number of individuals skilled in the use of these techniques who have shown interest in group data or patterns of interpersonal relations.

Northway and Wigdor (1947) have compared Rorschach responses of three groups of eighth grade boys and girls who were high, low, and middle in their sociometric status. Each group contained fifteen students. The reported findings indicate that the three groups were differentiated in response to the test, although it is rather difficult to evaluate these results in view of the dangers involved in quantitative treatment of untransformed Rorschach scores. Northway and Wigdor indicate that the high in status are characterized by "greater sensitivity in sensing the feelings of others, and a conscious striving for the approval of others" (p. 196). The middle group ". . . seems a more shallow, less introspective group with few anxiety or emotional disturbances" (p. 194). The low group "seems to be the most seriously disturbed. They show less ability to control their emotions, and seem to be a more egocentric, moody, and impulsive group" (p. 194). All of these interpretations are based upon differences between the groups on specific Rorschach scoring dimensions. In a further analysis the different status groups were compared in a general clinical interpretation of signs of disturbance. The low social choice group was divided into two subgroups, one made up of the recessives and the other the aggressive children. The findings indicated that there were more serious signs of disturbance in the recessive group than in any other. Pepinsky, Siegel, and Vanatta (1952), as part of an ingenious attempt to develop a criterion that could be used in evaluating therapeutic change, related sociometric choice to performance on the group Rorschach. Correlation of a large number of Rorschach variables with the number of sociometric choices given and received provided no evidence of any association between these measures. Multiple correlation analysis suggested the possibility of a relation between the number of choices given and two Rorschach scores (ΣM, Dd%).

A study by Lindzey and Goldwyn (1953) examined the relation between sociometric status, measured by choice and rejection, and certain P-F Study scores and TAT variables. A correlational comparison of a selected number of P-F Study variables and social status revealed very low correlations in the expected direction, some of which attained statistical significance. Further, when the ten individuals highest in social status were compared with the ten individuals lowest in social status, significant differences were observed on several of the P-F Study variables. The findings indicated that the high in social status were less extrapunitive, more intrapunitive, and showed lower group conformity ratings than the low in social status. TAT story ratings designed to measure anxiety and aggression did not relate to social status. In a study of student nurses Holzberg and Posner (1951) used a sociometric technique with criteria designed to secure information about the assertiveness of the group members. They found a significant positive relationship between sociometric status in this situation and scores on the A-S Reaction Study, as well as with ratings of assertiveness carried out by supervisors. There appeared to be no relation between sociometric scores and performance on the P-F Study or ratings of assertiveness derived from a modified TAT. Miller and Stine (1951) used a story-completion test in an effort to relate pregenital fixations to social choice as measured by a sociometric questionnaire. The expectation that individuals who were most characterized by pregenitality, as evidenced by the story-completion test, would be most rejected was, in general, verified. The investigators divided their subjects, primary school children, into five status groups in terms of the incidence of choice and rejection, and found a general increase in rated pregenitality as the degree of rejection increased. When the mean pregenital scores of the two high-status groups were combined and compared with the means of the low-status groups for each of eleven pregenital variables (orality, anality, passivity, direct aggression, etc.), the low-status group was higher on ten of eleven variables, five of these differences reaching the five percent level of significance. Thus, in adequately controlled circumstances, there appeared to be evidence that social choice was related to psychoanalytic variables as measured through the use of indirect or projective methods.

In general, we may suggest that the evidence is moderately convincing for the existence of regular relationships between certain sociometric dimensions and attributes measured by conventional personality instruments. Personality, at both overt and covert levels, appears to relate to sociometric variables in significant respects.

The purpose of this section has been to sample briefly some of the studies that have shown sociometric data to be associated with variables of psychological interest. In spite of serious limitations in the conceptions and procedures typical of many sociometric investigations, there appears to be evidence for substantial relations between aspects of behavior measured by sociometric techniques on the one hand and demographic, cultural, intellectual, attitude, and personality variables on the other.

SOCIOMETRIC MEASURES IN "ACTION"

Moreno's initial insistence that sociometric measures could be used properly only when the groups studied were restructured in the light of sociometric findings has led to a heavy tradition of "action" use of these measures. The earliest sociometric studies were carried out in settings where the findings were used to implement or modify practical programs of education or therapy. The ease with which the results of sociometric measures can be transformed into practical recommendations has become one of their most attractive features. Moreover, the use of sociometric findings to introduce change or restructure into the phenomena being studied does not preclude the likelihood of observing important empirical findings. In some cases this likelihood is actually enhanced.

Our discussion of sociometric applications is organized in terms of whether the application is primarily in the field of education, industry, military service, mental hospitals, or community service.

EDUCATIONAL APPLICATIONS

Sociometric measures first appeared in connection with the study of public school students and the inmates of a school for female delinquents (Moreno, 1934). The passage of time has seen continued application of these measures in educational systems and correctional institutions. Not only do schools present groups that are well designed to facilitate the use of sociometric measures, but also many of the problems faced by teachers and school administrators can be approached constructively through the use of sociometric measures. Thus, formation of cohesive, smooth-working groups is facilitated by information that the sociometric measure can provide. The task of sensitizing the teacher to the existence of interpersonal difficulties within her classes is facilitated by sociometric data, and efforts to combat these difficulties can be appraised through the use of sociometric measures. Further, and perhaps most important, these same techniques provide an excellent means of identifying the child whose relations with his classmates are sufficiently defective to suggest the likelihood of more serious difficulties in the future.

In general, sociometric measures may serve equally well the functions of improving normal interpersonal relations within the classroom and of diagnosing failures of interpersonal adjustment which may become the psychiatric casualties of a later period.

Probably the first reference to sociometric procedures in a classroom setting is an article by Moreno (1943b, originally published in 1932). In this paper he discussed the results of sociometric testing in a public elementary school, indicating the possibility of reassignment and also reporting surprising inaccuracy on the part of the teachers in judgments of the choice status of the students. The fallibility of the teacher as a judge of student relations indicates the importance of sociometric techniques as a supplementary measure. A number of additional studies (Bonney, 1947; Moreno, 1934) indicate that teachers' judgments, although variable in terms of the particular judge and the age of the students being evaluated, are not highly sensitive to individual differences in social acceptance. In particular, the teacher seems unable to discriminate effectively the object of main interest — the individual who receives a very low proportion of social choices. Perhaps the most extensive study of the accuracy of teachers' judgments of sociometric status was conducted by Gronlund (1950a, b) in forty sixth-grade classes. He found considerable variability among teachers in the accuracy of their judgments (Pearson correlations between teachers' ratings and sociometric ratings varied from .27 to .84) but most teachers were quite low in absolute accuracy. Further, teachers tended to overestimate the choice status of children they liked and to underestimate that of children they disliked.

The sensitivity of many teachers has been increased in a number of studies where the use of sociometric measures has supplemented judgments provided by teachers. Northway (1944) reports efforts made in cooperation with the school officials to bring about changes in the children's interpersonal relations in the direction of greater acceptance. An early use of sociometric measures to permit improving individual as well as group adjustment is reported by Kerstetter and Sargent (1940). In a ten-month study of a fifth grade class they observed the rise of a subgroup that developed deviant tendencies and produced marked conflict with both school authorities and the police. Following administration of an additional sociometric questionnaire, the members of the delinquent group were assigned new seats that separated them from each other, while at the same time taking advantage of their known preferences outside the group. The result of this "assignment therapy" appeared to be a trend toward dissolution of the delinquent group, as reflected in their responses to a third sociometric questionnaire administered three weeks later.

There have been numerous studies in which sociometric measures have been used to appraise the outcome of efforts to improve group or individual adjustment in classroom settings. In some cases the sociometric techniques were used simply to get at the outcome of the treatment, while in other cases they are involved in the treatment process itself. An attempt to improve social adjustment in a ninth grade class is reported by McClelland and Ratliff (1947), in which a sociometric questionnaire and a personality inventory were administered initially to appraise the adjustment of the class members. Following this, individuals who gave signs of particular adjustment difficulties were given positions that would place them in positive interaction with other members of the class. The class also embarked upon a self-improvement program and was divided into small groups, each of which had its own leader. A second sociometric questionnaire was administered which indicated a much more equitable distribution of choices, in spite of the fact that the same individuals were overchosen. The use of continuous sociometric testing in an educational setting, combined with "assignment therapy" and individual counseling for the poorly adjusted, is described by Shoobs (1947). He concludes that employment of these procedures over a six-month period resulted in both scholastic and behavioral improvement in the classroom.

Zeleny (1940b) reports a modified sociometric procedure that he feels may be used effectively to control the satisfactions and morale of individuals in a classroom setting. In a subsequent paper (Zeleny, 1941) he outlines a procedure for measuring social status in the classroom and discusses means whereby the social status of students can be improved. The results of a preliminary study of 25 students indicate that four weeks after these individuals had been exposed to individual guidance and sociometric regrouping there was a general increase in social status. Sociometric measures were used by Johnson (1939) to measure the effectiveness of attempts to bring about the integration of isolates in groups within a boys' club. The measures showed that in circumstances where the natural leader of the group was used as the agent to bring about the change, the isolate tended to be absorbed into the group.

The practice of regrouping individuals in a correctional institution is discussed by Moreno and Jennings (1944), and some indication is given of changes brought about by this regrouping. In particular, they point to the increased mutuality among the group members which occurred in a particular instance. They emphasize that one should not stop with regrouping the individual members, but rather that this should be combined with efforts at retraining key figures in the group. The use of "assignment therapy" in a classroom setting is described by Thrasher and Kerstetter (1947). Two college classes were given sociometric tests at the start of the term and a committee structure was established on this basis. At the end of the semester a second sociometric test was administered and comparison with the initial results showed a highly consistent increase for the group members in both choices received and choices given. Such findings are extremely difficult to interpret because of the absence of a control group. A study by Faunce and Beegle (1948) reports the use of sociometric measures to produce compatible work groups in a summer camp and also to measure the degree of integration in the camp.

It is characteristic of the studies we have discussed here that they involve sufficient weakness in pertinent empirical controls so that their results can be considered little more than suggestive. This does not mean that most of their findings may not be replicable, but simply that the evidence for firm generalizations in this area is not yet present.

We have mentioned here only a very small proportion of the studies that have dealt with sociometric measures in educational settings. However, these studies, in spite of their procedural limitations, are consistent in suggesting that there is a legitimate and important area of functioning for sociometric measures in the educative process.

INDUSTRIAL APPLICATIONS

The potential contribution of sociometric measurement in this area has increased greatly as psychologists in industrial settings have become less concerned with measurement of specific abilities and traits and have focused more and more upon describing the actual interpersonal setting in which jobs are carried out. The increasing interest in informal organization of working groups, as well as the frequent necessity of appraising morale, cohesiveness, relations to peer and authority figures, etc., makes it almost imperative that the industrial social scientist be familiar with sociometric measures.

One of the first applications of sociometric methods in industry was made by Jacobs (1945), who discusses the general value of such an approach and presents illustrations drawn from the use of these techniques in a business setting. He concludes that sociometric measures have value for measuring morale; selecting supervisors, foremen, and managers; determining informal groups; locating social factors that influence production; determining how well individuals are adjusting socially to their jobs; and measuring factors of cooperation, getting along with people, etc., for purposes of merit ratings and bonuses. Rogers (1946), in a discussion of human problems in industry, indicates her conviction that the work of Roethlisberger, Dickson, and others of the Harvard Business School group, and the contributions of Moreno, primarily in the form of sociometric measures, represent the two most important contributions to an understanding of the psychological problems of the worker in industry.

A study of 24 executives is reported by Browne (1951), who examined the relations existing among a small group of executives and also compared sociometric choice scores with various independent measures. Instead of asking the subjects to indicate the persons they would *like* to associate with, they were asked to indicate the persons they actually *did* associate with in the process of carrying out their jobs. The results are presented in the form of a sociogram that makes clear the formal organization of the company as well as the interaction patterns. The diagram shows that there are marked differences between divisions within the same company in the relations existing between executives. The frequency with which an individual was reported as involved in interactions with others showed some signs of positive correlation with questionnaire measures of responsibility, authority, and delegation of authority. Browne generalizes his experience by suggesting that the sociometric measure may be used in industry in order to study interpersonal relations, communication channels, and methods of performing leadership functions, as well as to suggest needed modifications in personnel relations. Weschler, Tannenbaum, and Talbot (1952) have discussed the value of a multidimensional sociometric approach to industrial problems.

The possibility that sociometric data may provide important information concerning the determinants and possible control of industrial accidents is suggested by two studies. Fuller and Baune (1951), studying third grade children, and Speroff and Kerr (1952), observing steel mill workers, both report that accident proneness appears to be negatively related to sociometric choice status. The latter investigators believe that the frequency of industrial accidents could be reduced through sociometric "reassignment" combined with active counseling of workers who are demonstrably low in social choice. The reader will find further discussion of the use of sociometric techniques in industrial settings in Chapter 22 and Chapter 29.

MILITARY APPLICATIONS

The increased activity of social scientists in military settings during the past decade has led to a number of studies in which sociometric measures have been employed in order to cope with problems of military importance. In general, the reasons that sociometric measures have been employed here are the same as those that have led to industrial applications — interest in morale, group cleavages, social adjustment, etc. An organization which asks individuals to function in group settings with a great deal of interdependency in their activities naturally is concerned with the quality of the interpersonal relations that characterize the groups, subgroups, and individuals.

Studies in this area may be conveniently divided into those that are primarily concerned with appraising the character of the individual's relations to others and those that are concerned with the quality of interpersonal relations for the group as a unit. One of the first reported studies using these techniques in an effort to provide information about the individual is reported by Sutherland and Fitzpatrick (1945). Their discussion of group techniques in the selection, training, therapy, and demobilization of the British army make clear that sociometric methods were used freely and felt to function effectively. An intensive sociometric investigation of 35 French army officers is described by Maucorps (1949), in which a number of correlates of sociometric indices were observed. He summarizes, "The sociometric procedure has proved to be an excellent selection tool. . . . Certain findings have revealed the practical importance of such an analysis as regards both the reducing of prejudice and the rational reconstruction of groups" (p. 79). Jennings (1949) discussed the use of sociometric and situation tests in military settings in Europe and the United States during the past war. She believes that the German use of these instruments provided less opportunity for freedom in response on the part of the subject, but that such techniques were used very widely on both continents.

An excellent study by Williams and Leavitt (1947) relates sociometric choice to performance in officers' candidate schools and, more importantly, to leadership performance in combat. The value of sociometric indices was examined through comparing the relative efficiency of sociometric choice by platoon members (with the criterion of "all-round ability as a combat officer") and a number of other indices, including self-ratings, leader ratings, personality inventory scores, and performance on objective ability tests. When these measures were related to success and failure in officers' candidate school and to ratings of performance as a leader in actual combat the sociometric measure showed a significant correlation with both criteria and *performed better than any of the other measures.* In discussing the superiority of the sociometric index the authors suggest: "Group members have more time to observe each other than do superior officers; they know each other in a realistic social context, and they react directly to each other's social-dominance behavior. All these are conditions favorable to informed judgment" (p. 291). A similar study is described by Baier (1947), in which a comparison of sociometric choice scores for a group of West Point cadets showed a correlation of .51 with ratings of efficiency provided by superior officers after the cadets had completed their training and spent eighteen months on active duty.

A study reported by Wherry and Fryer (1949), conducted in an Army officers' candidate school, investigated whether "buddy ratings" (modified sociometric techniques) measured leadership qualities rather than personal popularity. Two groups of officer candidates were studied. A large number of ratings by peers and superiors in addition to measures based on their performance during the training period and intelligence test scores were secured. All of the leadership ratings, in addition to academic grades, were submitted to a factor analysis which led to the identification of four factors: academic success, leadership, tactical standing, and group difference correction. Wherry and Fryer summarize:

> Buddy ratings appear to be the purest measure of "leadership." Tactical officers are also able to rate this trait but their ratings are quite heavily weighted by tactical standing. Academic instructors' ratings are practically useless for the evaluation of this trait." (p. 189)

Stogdill and Shartle (1948) and Stogdill (1949), in a study of naval officers, report the use of a sociometric technique that requires the subject to list his fellow officers in rank order in terms of the amount of time on the job spent interacting with them. The frequency with which the individual is mentioned by his fellow officers is positively correlated with his level in the staff organization and also shows some variation with the kind of administrative activity in which the individual spent most of his time.

Several studies have used sociometric techniques to describe or classify groups. A paper presented by Jenkins (1947) indicated the sensitivity of sociometric choice and rejection measures to morale differences between two combat naval air squadrons. When the two squadrons, known through observation to differ markedly in morale and efficiency, were compared in patterns of choice and rejection, pronounced and consistent differences were observable. The high morale squadron was characterized by positive choice of the commanding officer and executive officer, a high incidence of in-group positive choice, and an absence of cliques within the group. On the other hand, in the low morale squadron the commanding and executive officers tended to be either ignored or rejected; there was considerable within-group rejection, and there was evidence for the existence of distinct cliques. In an attempt to relate group cohesiveness to military performance, Goodacre (1951) used sociometric questions to provide an index of cohesiveness and also secured ratings of the effectiveness of performance for a dozen twelve-man squads while coping with field problems. His measure of cohesiveness was simply the number of choices made by the members of each group. The results indicated a significant positive association between group cohesiveness and effectiveness of squad performance in the field. Zeleny (1947b) reports the use of sociometric techniques to increase the compatibility of air force cadets who were assigned to the same team during their training. As the flying tasks required close coordination in the activities of pair-members, it seemed desirable to secure as much mutuality within the pairs as possible. When the interactions in the groups that had been established prior to the use of sociometric measures were examined, there was a relatively low incidence of mutuality. In general, Zeleny's study makes clear that it is possible through the use of sociometric measures and restructuring of groups to increase the incidence of positive choices and decrease the incidence of rejection. Although no direct evidence is presented to indicate that such a shift results in more effective performance, this expectation is justifiable in view of the tasks being performed by the groups and the findings of other investigators (Jenkins, 1947; Goodacre, 1951).

The studies we have reviewed provide evidence for the usefulness of sociometric techniques in military settings, especially if measurement of leadership, morale, social adjustment, and similar variables is of importance.

MENTAL HOSPITALS AND PSYCHOTHERAPY

While still in an exploratory stage, the application of sociometric procedures in mental hospital settings and in conjunction with the therapeutic process has already led to some suggestive findings. The importance of interpersonal relations in personal adjustment, in addition to their crucial status in the treatment process, makes it understandable that sociometric measures should have some function in these problem areas.

Jenkins and Curran (1940) report the use of a sociometric questionnaire to secure information concerning the social relations existing in a group of about 40 adolescents in a psychiatric observation ward. The investigators report a number of illustrative sociograms and discuss them in relation to the known attributes of some of the patients included in the sociograms. A more recent study by Hyde and York (1948) describes a time-sampling technique that can be used in the observation of ward behavior and which leads to data that can be presented in the form of a sociogram. The investigators suggest that the device leads to a number of uses, including aid in individual diagnosis, selection of patients for group psychotherapy, providing evidence of the efficiency of patients' interpersonal relations (including signs of improvement or recovery), and providing relevant information for hospital personnel that may motivate them to perform more effectively. Sharing the belief of these earlier investigators that sociometric measures provide empirical control over an important aspect of individual adjustment, Pepinsky *et al.* (1952) attempted to develop a modified sociometric questionnaire that could be used as a criterion against which therapeutic efforts could be tested. A less elaborate attempt to develop a measure of social adjustment for a psychiatric setting was made by Schauer (1946), who used a simple sociometric technique in a small mental hospital population. His results show a very low amount of reciprocity in choice among patients, but he made no detailed attempt to determine how successful the sociometric measure was as an index of adjustment.

COMMUNITY SERVICE

The economy and speed with which sociometric measures can provide information concerning the personal ties and repulsions within relatively large groups have made them an efficient tool in the hands of individuals concerned with studying and facilitating various community functions. Thus, the social worker, community leader, and resettlement agent have all found these measures valuable.

The use of sociometric findings in the study of open communities was perhaps first reported by Wolman (1937), who studied a resettlement community of 35 families, composed of individuals drawn from low-income brackets in urban areas. He administered two sociometric tests six months apart. The first administration was at the start of the project and the results were used in determining the location of the families. The second administration was designed to discover what changes the six months of community life had brought about. Although the author discusses the changes in social position of the families individually, the only general change reported was that two of the six families originally unchosen had become integrated into the group. Closely related to this study are a series of investigations by Loomis carried out under the auspices of the U. S. Department of Agriculture. Loomis and Davidson (1939a), in a study of a rural resettlement community, utilized a modified sociometric technique, asking the identity of persons who actually visited the family being interviewed. They conclude: ". . . it is most important that the chain of informal social relationships on the projects should be studied. The present study conclusively demonstrates that such relationships are important. The families which remained and the families which moved away from the project . . . tended to constitute in-groups" (p. 78). In a second study, Loomis and Davidson (1939b) used the same technique to examine the extent to which two years of community existence had served to increase group integration. When the sociometric findings were compared with participant observer reports of cleavages present earlier, they indicated signs of increased community integration. Loomis (1941) used the same method in a Spanish-American village with the intent of arriving at insights that would be of value to Federal agencies concerned with resettling similar groups. He concluded that, to function smoothly, any attempt to resettle such groups must take into consideration the extended family. A more recent study by Infield (1947) reports the use of sociometric measures in a study of a Canadian cooperative community.

The utility of sociometric measurement in the hands of the social worker has been the subject of considerable discussion (Deutschberger, 1950; Green, 1950; Moreno, 1950). A brief paper by Kerstetter (1943) has outlined the uses of sociometric tools for social agencies.

A study is reported by Cologne (1943) in which members of a self-help community center were reassigned to work groups on the basis of sociometric data after the observation of hostility and tension among the members of the center. She demonstrated that the reassignment could be made in such a way as to satisfy, to some extent, the choices of a very large proportion of the group members.

The use of sociometric devices to detect patterns of influence within given communities and to aid in the selection of community leaders has been a topic of considerable interest. The application of sociometric measures in selecting "block leaders" as well as in assisting with the problem of how a "block" should be defined is described by Sanders (1949). Many community functions necessitate the use of some small organized unit such as the block in order to preserve channels of communication. In spite of a general tendency to define a block as an area bounded on four sides by streets, Sanders' sociometric data, collected in two different socioeconomic areas, indicates that most of the neighborhood interpersonal relations are between individuals on opposite sides of streets. He concludes that it is generally more fruitful to define a block in this way. An advantage of this approach is that it makes it possible to estimate who in the group is not likely to be contacted effectively by any particular leader. Several papers by Stewart (1947a, b) indicate the feasibility of modified sociometric techniques as a means of determining the influential individuals in a sizeable community.

The investigations we have discussed make clear that sociometric measures can serve a number of valuable functions for the student or practitioner working in community settings. The further development of large-scale sociometric measurement may be expected to enhance this value.

NEGLECTED CONSIDERATIONS AND LIMITATIONS IN THE USE OF SOCIOMETRIC MEASURES

In our discussion thus far we have intentionally presented sociometric measures in a favorable light. Our purpose has been to indicate those areas in which such measures are useful and the kinds of problems and concepts with which they interact fruitfully. At this point, we should like to provide a partial antidote for this consistent optimism in the form of a small number of cautions or limitations with which the user of sociometric techniques should be familiar. In some cases these cautions refer to the manner in which investigators have characteristically employed these techniques, while in other cases the fault is more general and inheres in the very nature of the sociometric technique.

EXCESSIVE DEPENDENCE UPON SOCIOMETRIC DATA

Sociometric measures provide only one of many possible means of viewing interpersonal relations. Knowledge of the formal or official organization of a group may be fully as important as sociometric data. Further, the evidence that can be provided through systematic observation of the group is in many cases essential. An adequate specification of the physical, cultural, and sociological setting, as well as a detailed psychological appraisal of the individuals involved, is often a prerequisite to full understanding of the phenomena to which sociometric techniques are customarily applied.

The inadvisability of securing sociometric data alone has frequently been noted by writers in this area. Moreno, in *Who Shall Survive?*, advocates the use of spontaneity tests (role playing) and interviews in addition to the sociometric test, in order to give information about the determinants of the sociometric responses. The ineffectiveness of this and subsequent appeals is in part attributable to the tendency of these writers to represent sociometric data as more "real," "spontaneous," or "basic" than other data that can be collected in connection with the same events. The desire to give sociometric data a special empirical status, subordinating and ignoring other kinds of data, may diminish the motivation to introduce the variety of other measures necessary for the most fruitful application of sociometric measures.

Whatever the determinants of this constrictive tendency may have been, it is clear that sociometric research has had more than its share of investigations where internal relations are sought. The examining of one aspect of sociometric response in relation to another, or the relating of sociometric response to the most obvious independent measures (age, sex, religion), can make only a limited contribution once the early phases of investigation are past.

SCARCITY OF SYSTEMATIC INVESTIGATION

It is quite characteristic of social scientists to let caprice and convenience determine both the problem to be explored and the manner in which it is approached. This tendency, however, seems especially marked among persons doing sociometric research. There has been very little in the way of systematic, theoretically derived research aimed at exploring in a regular fashion the kinds of conditions and variables that are related to sociometric response. Most of the evidence accumulated in this area has resulted from the use of convenient or interesting measures, with no attempt to explore those relations that might have been predicted in advance, have greatest theoretical relevance, or that would be most likely to stimulate further investigation.

The relative scarcity of systematic, cumulative investigations is intimately related to the tendency to rely heavily upon sociometric measures alone. The person who permits his empirical problem to be determined by casual interest and the availability of data is very likely to be content with sociometric data alone, while the individual who pursues a problem derived from theory or which is part of an orderly plan of research is almost certain to introduce other measures. Related to the lack of systematic direction of research has been the unimaginativeness with which studies have been carried out repeatedly with educational and institutional groups, while very little attention has been paid to the many other pertinent groups within our society. This lack of concern for representativeness in the subjects studied is perhaps defensible when the investigator is concerned with a general empirical relationship, but when he is simply concerned with taxonomy or description of group structure, he must be concerned with the degree to which his subjects represent a definable population.

It seems quite likely that the value of studies that make casual and unsystematic use of sociometric measures is steadily decreasing. Explicit and multivariate conceptions of the nature of interpersonal relations can be expected to lead to research that will be both cumulative and of more basic importance in the understanding of group process. Further, a more adequate understanding of sociometric instruments is likely to accrue from such systematic programs of research.

The facility with which Moreno has contrived and presented speculative theoretical constructions in connection with his empirical findings, and his lack of concern for a rigorous linking of these formulations to empirical data may have served as a deterrent to a more cautious and persistent approach to the relation between theory and sociometric data.

An interesting step has been made by Leeman (1952) in a direction that may eventually lead to a more systematic approach to sociometric data. He presents a base-line mathematical model to represent sociometric choice. In mathematical models the initial assumptions are, of course, crucial to the subsequent derivations, and in the "base-line model" emphasis is placed upon choosing assumptions that can be handled rather than those that are necessarily most accurate. The correctness of the model's predictions concerning the empirical world is presumably an outcome of subsequent revision. Leeman presents derivations from his model for three- and four-man groups and confronts his derivations with empirical data. The results of this comparison are promising. There are many difficulties involved in such an approach — Leeman indicates the practical limit of his model as the four-man group and further points to the complexities involved in making derivations from the model. Nevertheless, this direction represents one way in which more efficient data collection and interpretation might evolve.

INSUFFICIENT ATTENTION TO THE SELECTION OF THE SOCIOMETRIC CRITERION

All too often sociometric investigators have used much of their time and energy in the collection of data, and have given little thought to the criterion of the sociometric question. The importance of selecting an appropriate criterion has been enunciated frequently but seems to have had only a limited audience. Moreno (1937a) distinguished between the *essential* and the *auxiliary* criteria — the essential criteria are those that refer to the activity which is at the center of the investigator's interest. Quite frequently investigators have used auxiliary criteria to define or measure variables such as leadership, social status, etc., with the consequence that they secured contaminated measures and results that were difficult to interpret.

Relevant to this issue is the distinction proposed by Criswell (1949) betwen one-way and two-way sociometric questions. As she points out, certain sociometric criteria (sit next to, room with) refer to relationships that are reciprocal, where the extent to which choices are mutual or shared is an all-important question. Other criteria (lead a squad, represent a school at a conference) refer to relationships where mutual choice is relatively unimportant. The crucial question is how frequently the person is chosen, not whether he is chosen by those whom he chooses. A further relevant distinction is that between *psyche-group* and *socio-group* (Jennings, 1950). These distinctions are important in the present discussion because they point to the fact that if the criterion used is not selected carefully, the investigator may well secure data that do not lead naturally to the kinds of analyses in which he is interested.

Not only is it important that investigators consider carefully the question of what criterion is most appropriate for their special use, it is also important that future investigations aim directly at the question of what choices under various criteria mean (relate to), or the kinds of variables that can best be measured through the use of particular criteria.

One of the most frequent objections raised to sociometric measures is the claim that the investigator does not know what a choice "means" for each individual. Thus, for one subject choice is evidence of personal affection, for another it is an intellectual evaluation, while for a third it is a mechanical answer conforming to the experimenter's demands. This objection may refer to the general question of the relevance of various criteria for providing data pertinent to particular dimensions. Viewed in this light it is legitimate, or at least points to an important area of inquiry. However, it is not proper to ask that the investigator know what a given response "actually means" to every respondent. Such a demand, if taken literally, would rule out virtually every data collection procedure commonly used in the social sciences. It is no more necessary that we know what a given sociometric response means uniquely to each subject than that we know what each check on a questionnaire or each statement in an interview means. The important question is the extent to which an order imposed upon subjects from this kind of response data can be related to pertinent independent measures.

INADEQUATE QUANTITATIVE TREATMENT OF DATA

In spite of the tremendous forward steps that have been made in the analysis of sociometric data, characteristic flaws remain in many studies. Most serious of these is the tendency to capitalize illegitimately upon chance variation and thus to treat random variations as representing significant findings. Frequently this is accomplished by conducting a large number of tests of significance and then ignoring those relations that fail to materialize, while concentrating upon the results that achieve respectable significance. This leads to a false sense of security and the acceptance of relations as demonstrated that are unlikely to prove replicable. Even this is more defensible than the practice of reporting numerical results and then discussing observable differences with no attempt to assess the role of chance. Fortunately, the incidence of such studies has shown a marked decrease in recent years. There are many more sophisticated problems encountered in the quantitative analysis of sociometric data and some of these have been dealt with in previous sections of this chapter.

Summary

We have attempted to outline the ways in which sociometric data can be collected, appropriate methods for analyzing these data, and some of the areas of investigation and application where sociometric measures may be useful. Although we have noted certain limitations in sociometric measures and their customary application, the general impression we have tried to communicate is that the measures are of substantial value to the investigator concerned with group processes or the individual in a group setting. Further, the value of these measures is certain to be enhanced through the development of more sophisticated techniques for the quantitative analysis of sociometric data and through using these measures more extensively in conjunction with other methods of data collection.

References

Ashley-Montagu, M. F. Sociometric methods in anthropology. *Sociometry,* 1945, *8,* 62-63.

Baier, D. E. Note on "A review of leadership studies with particular reference to military problems." *Psychol. Bull.,* 1947, *44,* 466-467.

Bain, R. Sociometry and social measurement. *Sociometry,* 1943, *6,* 206-213.

Barker, R. G. The social interrelations of strangers and acquaintances. *Sociometry,* 1942, *5,* 169-179.

Baron, D. Personal-social characteristics and classroom social status: A sociometric study of fifth- and sixth-grade girls. *Sociometry,* 1951, *14,* 32-42.

Bassett, R. E. Sampling problems in influence studies. *Sociometry,* 1948, *11,* 320-328.

Bell, G. B., & French, R. L. Consistency of individual leadership position in small groups of varying membership. *J. abnorm. soc. Psychol.,* 1950, *45,* 764-767.

Beum, C. O., & Brundage, E. G. A method for analyzing the sociomatrix. *Sociometry,* 1950, *13,* 141-145.

Beum, C. O., & Criswell, Joan H. Application of machine tabulation methods to sociometric data. *Sociometry,* 1947, *10,* 227-232.

Bjerstedt, A. A "chess-board sociogram" for sociographic representation of choice directions and for the analysis of "sociometric locomotions." *Sociometry,* 1952, *15,* 244-262.

Bock, R. D., & Husain, Suraya Z. An adaptation of Holzinger's B-coefficients for the analysis of sociometric data. *Sociometry,* 1950, *13,* 146-153.

Bonney, M. E. A study of social status on the second-grade level. *J. genet. Psychol.,* 1942, *60,* 271-305.

Bonney, M. E. Personality traits of socially successful and socially unsuccessful children. *J. educ. Psychol.,* 1943a, *34,* 449-472.

Bonney, M. E. The constancy of sociometric scores and their relationship to teacher judgments of social success, and to personality self-ratings. *Sociometry,* 1943b, *6,* 409-424.

Bonney, M. E. Relationships between social success, family size, socioeconomic home background, and intelligence among school children in grades III to V. *Sociometry,* 1944, *7,* 26-39.

Bonney, M. E. A sociometric study of the relationship of some factors to mutual friendships on the elementary, secondary, and college levels. *Sociometry,* 1946, *9,* 21-47.

Bonney, M. E. Sociometric study of agreement between teacher judgments and student choices: In regard to the number of friends possessed by high school students. *Sociometry,* 1947, *10,* 133-146.

Bonney, M. E. A study of friendship choices in college in relation to church affiliation, in-church preferences, family size, and length of enrollment in college. *J. soc. Psychol.,* 1949, *29,* 153-166.

Borgatta, E. F. A diagnostic note on the construction of sociograms and action diagrams. *Group Psychotherapy,* 1951, *3,* 300-308.

Bronfenbrenner, U. A constant frame of reference for sociometric research. *Sociometry,* 1943, *6,* 363-397.

Bronfenbrenner, U. A constant frame of reference for sociometric research: Part II. Experiment and inference. *Sociometry,* 1944a, *7,* 40-75.

Bronfenbrenner, U. The graphic presentation of sociometric data. *Sociometry,* 1944b, *7,* 283-289.

Browne, C. G. Study of executive leadership in business. IV. Sociometric pattern. *J. appl. Psychol.,* 1951, *35,* 34-37.

Byrd, E. A study of validity and constancy of scores in a sociometric test. *Sociometry,* 1951, *14,* 175-181.

Cervinka, V. A dimensional theory of groups. *Sociometry,* 1948, *11,* 100-107.

Chabot, J. A simplified example of the use of matrix multiplication for the analysis of sociometric data. *Sociometry,* 1950, *13,* 131-140.

Chapin, F. S. Trends in sociometrics and critique. *Sociometry,* 1940, *3,* 245-262.

Chapin, F. S. The relation of sociometry to planning in an expanding social universe. *Sociometry,* 1943, *6,* 234-240.

Chapin, F. S. Sociometric stars as isolates. *Amer. J. Soc.*, 1950, *56*, 263-267.

Chapin, F. S. Comment. *Amer. J. Soc.*, 1951, *57*, 264.

Cologne, Rose Experimentation with sociometric procedure in a self-help community center. *Sociometry*, 1943, *6*, 27-67.

Criswell, Joan H. Racial cleavage in Negro-white groups. *Sociometry*, 1937, *1*, 81-89.

Criswell, Joan H. A sociometric study of race cleavage in the classroom. *Arch. Psychol.*, No. 235, 1939.

Criswell, Joan H. The saturation point as a sociometric concept. *Sociometry*, 1942, *5*, 146-150.

Criswell, Joan H. Sociometric methods of measuring group preferences. *Sociometry*, 1943, *6*, 398-408.

Criswell, Joan H. Foundations of sociometric measurement. *Sociometry*, 1946, *9*, 7-13.

Criswell, Joan H. The measurement of group integration. *Sociometry*, 1947, *10*, 259-267.

Criswell, Joan H. Sociometric concepts in personnel administration. *Sociometry*, 1949, *12*, 287-300.

Criswell, Joan H. Notes on the constant frame of reference problem. *Sociometry*, 1950, *13*, 93-107.

Criswell, Joan H., & Jennings, Helen H. A critique of Chapin's "sociometric stars as isolates." *Amer. J. Soc.*, 1951, *57*, 260-264.

Danielsson, B. Some attraction and repulsion patterns among Jibaro Indians. *Sociometry*, 1949, *12*, 83-105.

Deutschberger, P. Interaction patterns in changing neighborhoods: New York and Pittsburgh. *Sociometry*, 1946, *9*, 303-315.

Deutschberger, P. Sociometry and social work. *Sociometry*, 1950, *13*, 8-21.

Edwards, Daisy S. The constant frame of reference problem in sociometry. *Sociometry*, 1948, *11*, 372-379.

Eng, E., & French, R. L. The determination of sociometric status. *Sociometry*, 1948, *11*, 368-371.

Faunce, D., & Beegle, J. A. Cleavages in a relatively homogeneous group of rural youth: An experiment in the use of sociometry in attaining and measuring integration. *Sociometry*, 1948, *11*, 207-216.

Festinger, L. The analysis of sociograms using matrix algebra. *Hum. Rel.*, 1949, *2*, 153-158.

Festinger, L., Schachter, S., & Back, K. *Social pressures in informal groups: A study of human factors in housing.* New York: Harpers, 1950.

Flemming, G. A factor analysis of the personality of high school leaders. *J. appl. Psychol.*, 1935, *19*, 597-605.

Foa, U. G. Scale and intensity analysis in sociometric research. *Sociometry*, 1950, *13*, 358-362.

Forsyth, Elaine, & Katz, L. A matrix approach to the analysis of sociometric data: preliminary report. *Sociometry*, 1946, *9*, 340-347.

French, R. L. Sociometric status and individual adjustment among naval recruits. *J. abnorm. soc. Psychol.*, 1951, *46*, 64-72.

French, R. L., & Mensh, I. N. Some relationships between interpersonal judgments and sociometric status in a college group. *Sociometry*, 1948, *11*, 335-345.

Fuller, Elizabeth M., & Baune, Helen B. Injury-proneness and adjustment in a second grade: A sociometric study. *Sociometry*, 1951, *14*, 210-225.

Gibb, C. A. The sociometry of leadership in temporary groups. *Sociometry*, 1950, *13*, 226-243.

Goodacre, D. M., III The use of a sociometric test as a predictor of combat unit effectiveness. *Sociometry*, 1951, *14*, 148-152.

Goodnow, R. E., & Tagiuri, R. Religious ethnocentrism and its recognition among adolescent boys. *J. abnorm. soc. Psycol.*, 1952, *47*, 316-320.

Gottheil, E. Sociometric technique and experimental method in social psychology. *J. soc. Psychol.*, 1952, *35*, 9-21.

Green, Helen D. Sociometry and social intergroup work. *Sociometry*, 1950, *13*, 22-28.

Gronlund, N. E. The accuracy of teachers' judgments concerning the sociometric status of sixth-grade pupils: Part I. *Sociometry*, 1950a, *13*, 197-225.

Gronlund, N. E. The accuracy of teachers' judgments concerning the sociometric status of sixth-grade pupils: Part II. *Sociometry*, 1950b, *13*, 329-357.

Grossmann, Beverly, & Wrighter, Joyce. The relationship between selection-rejection and intelligence, social status, and personality amongst sixth-grade children. *Sociometry*, 1948, *11*, 346-355.

Hartshorne, H., May, M. A., & Maller, J. B. *Studies in the nature of character: II. Studies in service and self-control.* New York: Macmillan, 1929.

Holzberg, J. D., & Posner, Rita The relationship of extrapunitiveness on the Rosenzweig Picture-Frustration Study to aggression in overt behavior and fantasy. *Amer. J. Orthopsychiat.*, 1951, *21*, 767-779.

Hunt, J. McV., & Solomon, R. L. The stability and some correlates of group-status in a summer-camp group of young boys. *Amer. J. Psychol.*, 1942, *55*, 33-45.

Hyde, R. W., & York, R. H. A technique for investigating interpersonal relationships in a mental hospital. *J. abnorm. soc. Psychol.*, 1948, *43*, 287-299.

Infield, H. F. A veterans' cooperative land settlement and its sociometric structure. *Sociometry*, 1947, *10*, 50-70.

Jacobs, J. H. The application of sociometry to industry. *Sociometry*, 1945, *8*, 181-198.

Jenkins, J. G. The nominating technique, its uses and limitations. *Amer. J. Psychol.*, 1947, *2*, 433. (Abstract)

Jenkins, R. L., & Curran, F. J. The evolution and persistence of groups in a psychiatric observation ward. *J. soc. Psychol.*, 1940, *12*, 279-289.

Jennings, Helen H. Structure of leadership — development and sphere of influence. *Sociometry*, 1937, *1*, 99-143.

Jennings, Helen H. *Leadership and isolation.* New York: Longmans, Green, 1943.

Jennings, Helen H. Leadership and sociometric choice. *Sociometry*, 1947, *10*, 32-49.

Jennings, Helen H. Military use of sociometric and situation tests in Great Britain, France, Germany and the United States. *Sociometry*, 1949, *12*, 191-201.

Jennings, Helen H. *Leadership and isolation.* 2nd ed. New York: Longmans, Green, 1950.

Johnson, A. D. An attempt at change in interpersonal relationships. *Sociometry*, 1939, *2*, 43-48.

Katz, L. On the matrix analysis of sociometric data. *Sociometry*, 1947, *10*, 233-241.

Katz, L. Punched card technique for the analysis of multiple level sociometric data. *Sociometry*, 1950, *13*, 108-122.

Katz, L. A new status index derived from sociometric analysis. *Psychometrika*, 1953, *18*, 39-43.

Kerstetter, Leona M. Sociometry and social agencies. *Sociometry*, 1943, *6*, 255-257.

Kerstetter, Leona M., & Sargent, J. Re-assignment therapy in the classroom. *Sociometry*, 1940, *3*, 292-306.

Kuhlen, R. G., & Bretsch, H. S. Sociometric status and personal problems of adolescents. *Sociometry*, 1947, *10*, 122-132.

Leeman, C. P. Patterns of sociometric choice in small groups: A mathematical model and related experimentation. *Sociometry*, 1952, *15*, 220-243.

Lemann, T. B., & Solomon, R. L. Group characteristics as revealed in sociometric patterns and personality ratings. *Sociometry*, 1952, *15*, 7-90.

Lindzey, G., & Goldwyn, R. Validity of the P-F Study. *J. Pers.*, 1954, *22*, 519-547.

Lindzey, G., & Urdan, J. A. Personality and social choice. *Sociometry*, 1954, *17*, 47-63.

Longmore, T. W. A matrix approach to the analysis of rank and status in a community in Peru. *Sociometry*, 1948, *11*, 192-206.

Loomis, C. P. Informal groupings in a Spanish-American village. *Sociometry*, 1941, *4*, 36-51.

Loomis, C. P. Ethnic cleavages in the southwest as reflected in two high schools. *Sociometry*, 1943, *6*, 7-26.

Loomis, C. P. Political and occupational cleavages in a Hanoverian village, Germany. *Sociometry*, 1946, *9*, 316-333.

Loomis, C. P. The most frequently chosen sociogram: or The seduction of rural sociologists by the neighborhood theory. *Sociometry*, 1948, *11*, 230-234.

Loomis, C. P., Baker, W. B., & Proctor, C. The size of the family as related to social success of children. *Sociometry*, 1949, *12*, 313-320.

Loomis, C. P., Beegle, J. A., & Longmore, T. W. Critique of class as related to social stratification. *Sociometry*, 1947, *10*, 319-337.

Loomis, C. P., & Davidson, D. M., Jr. Measurement of the dissolution of in-groups in the integration of a rural resettlement project. *Sociometry*, 1939a, *2*, 84-94.

Loomis, C. P., & Davidson, D. M., Jr. Sociometrics and the study of new rural communities. *Sociometry*, 1939b, *2*, 56-76.

Loomis, C. P., & Pepinsky, H. B. Sociometry, 1937-1947: Theory and methods. *Sociometry*, 1948, *11*, 262-286.

Loomis, C. P., & Powell, R. M. Sociometric analysis of class status in rural Costa Rica — A peasant community compared with an hacienda community. *Sociometry*, 1949, *12*, 144-157.

Loomis, C. P., & Proctor, C. The relationship between choice status and economic status in social systems. *Sociometry*, 1950, *13*, 307-313.

Luce, D. R. Connectivity and generalized cliques in sociometric group structure. *Psychometrika*, 1950, *15*, 169-190.

Luce, D. R., & Perry, A. D. A method of matrix analysis of group structure. *Psychometrika*, 1949, *14*, 95-116.

Lundberg, G. A. Social attraction patterns in a rural village: A preliminary report. *Sociometry*, 1937, *1*, 77-80.

Lundberg, G. A. Discussion of sociometry. *Sociometry*, 1943, *6*, 219-220.

Lundberg, G. A., & Beazley, Virginia "Consciousness of kind" in a college population. *Sociometry*, 1948, *11*, 59-74.

Lundberg, G. A., & Dickson, Lenore Inter-ethnic relations in a high-school population. *Amer. J. Soc.*, 1952, *58*, 1-10.

Lundberg, G. A., Hertzler, Virginia B., & Dickson, Lenore. Attraction patterns in a university. *Sociometry*, 1949, *12*, 158-169.

Lundberg, G. A., & Lawsing, Margaret The sociography of some community relations. *Amer. soc. Rev.*, 1937, *2*, 318-335.

Lundberg, G. A., & Steele, Mary Social attraction-patterns in a village. *Sociometry*, 1938, *1*, 375-419.

McClelland, F. M., & Ratliff, J. A. The use of sociometry as an aid in promoting social adjustment in a ninth-grade home-room. *Sociometry*, 1947, *10*, 147-153.

McKinney, J. C. An educational application of a two-dimensional sociometric test. *Sociometry*, 1948, *11*, 356-367.

Massarik, F., Tannenbaum, R., Kahane, M., & Weschler, I. R. Sociometric choice and organizational effectiveness. *Sociometry*, 1953, *16*, 211-238.

Maucorps, P. H. A sociometric inquiry in the French Army. *Sociometry*, 1949, *12*, 46-80.

Mead, Margaret Some relations between cultural anthropology and sociometry. *Sociometry*, 1947, *10*, 312-318.

Miller, D. R., & Stine, Margaret E. The prediction of social acceptance by means of psychoanalytic concepts. *J. Pers.*, 1951, *20*, 162-174.

Moreno, J. L. *Who shall survive?* Washington, D. C.: Nervous and Mental Disease Monograph, No. 58, 1934.

Moreno, J. L. Sociometry in relation to other social sciences. *Sociometry*, 1937, *1*, 206-219.

Moreno, J. L. Sociometry in action. *Sociometry*, 1942, *5*, 298-315.

Moreno, J. L. Sociometry and the cultural order. *Sociometry*, 1943a, *6*, 299-344.

Moreno, J. L. Sociometry in the classroom. *Sociometry*, 1943b, *6*, 425-428.

Moreno, J. L. Sociogram and sociomatrix. *Sociometry*. 1946, *9*, 348-349.

Moreno, J. L. Note on cohesion in social groups. *Sociometry*, 1950a, *13*, 176.

Moreno, J. L. The sociometric approach to social case work. *Sociometry*, 1950b, *13*, 172-175.

Moreno, J. L., & Jennings, Helen H. Statistics of social configurations. *Sociometry*, 1938, *1*, 342-374.

Moreno, J. L., & Jennings, Helen H. Sociometric methods of grouping and regrouping: With reference to authoritative and democratic methods of grouping. *Sociometry*, 1944, *7*, 397-414.

Moreno, J. L., Jennings, Helen H., & Sargent, J. Time as a quantitative index of interpersonal relations. *Sociometry*, 1940, *3*, 62-80.

Newcomb, T. M. *Personality and social change.* New York: Dryden, 1943.

Newstetter, W. I., Feldstein, M. J., & Newcomb, T. M. *Group adjustment — A study in experimental sociology.* School of Applied Social Sciences, Western Reserve Univ., Cleveland, 1938.

Northway, Mary L. A method for depicting social relationships obtained by sociometric testing. *Sociometry*, 1940, *3*, 144-150.

Northway, Mary L. Outsiders: A study of the personality patterns of children least acceptable to their age mates. *Sociometry*, 1944, *7*, 10-25.

Northway, Mary L. Sociometry and some challenging problems of social relationships. *Sociometry*, 1946, *9*, 187-198.

Northway, Mary L. *A primer of sociometry.* Toronto: University of Toronto Press, 1952.

Northway, Mary L., & Quarrington, B. Depicting inter-cultural relations. *Sociometry*, 1946, *9*, 334-339.

Northway, Mary L., & Wigdor, Blossom T. Rorschach patterns related to the sociometric status of school children. *Sociometry*, 1947, *10*, 186-199.

Pepinsky, H. B., Clyde, R. J., Olesen, Barbara A., & Vanatta, E. L. The criterion in counseling: I. Individual personality and behavior in a social group. *Educ. Psychol. Measmt.*, 1952, *12*, 178-193.

Pepinsky, H. B., Siegel, L., & Vanatta, E. L. The criterion in counseling: A group participation scale. *J. abnorm. soc. Psychol.*, 1952, *47*, 415-419.

Pepinsky, Pauline N. The meaning of "validity" and "reliability" as applied to sociometric tests. *Educ. Psychol. Measmt.*, 1949, *9*, 39-49.

Pintner, R., Forlano, G., & Freedman, H. Personality and attitudinal similarity among classroom friends. *J. appl. Psychol.*, 1937, *21*, 48-65.

Polansky, N., Lippitt, R., & Redl, F. The use of near-sociometric data in research on group treatment processes. *Sociometry*, 1950, *13*, 39-62.

Powell, R. M. A comparative social class analysis of San Juan Sur, and Attiro, Costa Rica. *Sociometry*, 1951, *14*, 182-202.

Proctor, C. H. *Turrialba, Costa Rica — Comparison of social system and change in small farm and hacienda villages.* East Lansing: Michigan State College Press, in press.

Proctor, C. H., & Loomis, C. P. Analysis of sociometric data. In Marie Jahoda, M. Deutsch, & S. W. Cook (Eds.), *Research methods in social relations: With especial reference to prejudice.* Part Two. New York: Dryden, 1951. Pp. 561-585.

Richardson, Helen M. Community of values as a factor in friendships of college and adult women. *J. soc. Psychol.*, 1940, *11*, 303-312.

Richmond, Winifred Sociometric tests in a training school for nurses. *Sociometry*, 1950, *13*, 29-38. (Originally published in *Sociometric Review*, 1936, 41-49.)

Rogers, Maria Problems of human relations in industry. *Sociometry*, 1946, *9*, 350-371.

Ross, I. C., & Harary, F. On the determination of redundancies in sociometric chains. *Psychometrika*, 1952, *17*, 195-208.

Sanders, I. T. The use of block leaders in effective community mobilization. *Sociometry*, 1949, *12*, 265-275.

Sanderson, D. Discussion of sociometry. *Sociometry*, 1943, *6*, 214-218.

Schauer, G. Social adjustment in a mental hospital community. *Sociometry*, 1946, *9*, 144.

Seagoe, Mary V. Factors influencing the selection of associates. *J. educ. Res.*, 1933, *27*, 32-40.

Seeman, M. A situational approach to intra-group Negro attitudes. *Sociometry*, 1946, *9*, 199-206.

Shoobs, N. E. Sociometry in the classroom. *Sociometry*, 1947, *10*, 154-164.

Smelser, N. Sociometric structure and its correlates: Social class background and courtship-marriage attitudes. Unpublished honors thesis, Harvard University, 1952.

Smith, M. Some factors in friendship selections of high school students. *Sociometry*, 1944, *7*, 303-310.

Smucker, O. Measurement of group tension through the use of negative sociometric data. *Sociometry*, 1947, *10*, 376-383.

Smucker, O. Near-sociometric analysis as a basis for guidance. *Sociometry*, 1949, *12*, 326-340.

SPEROFF, B., & KERR, W. Steel mill "hot strip" accidents and interpersonal desirability values. *J. clin. Psychol.*, 1952, *8*, 89-91.

STEWART, F. A. Sociometric testing at the adult level. *Sociometry*, 1946, *9*, 147-148.

STEWART, F. A. A sociometric study of influence in Southtown. *Sociometry*, 1947a, *10*, 11-31.

STEWART, F. A. A study of influence in Southtown: II. *Sociometry*, 1947b, *10*, 273-286.

STEWART, F. A. Some sampling problems in sociometric surveys. *Sociometry*, 1948, *11*, 301-307.

STEWART, ISABEL A. An interviewer's report on adult sociometric study. *Sociometry*, 1948, *11*, 308-319.

STOGDILL, R. M. The sociometry of working relationships in formal organizations. *Sociometry*, 1949, *12*, 276-286.

STOGDILL, R. M., & SHARTLE, C. L. Methods for determining patterns of leadership behavior in relation to organization structure and objectives. *J. appl. Psychol.*, 1948, *32*, 286-291.

SUTHERLAND, J. D., & FITZPATRICK, G. A. Some approaches to group problems in the British Army. *Sociometry*, 1945, *8*, 443-455.

TAGIURI, R. Relational analysis: An extension of sociometric method with emphasis upon social perception. *Sociometry*, 1952, *15*, 91-104.

THRASHER, F. M., & KERSTETTER, LEONA M. Sociometry and an activity program on the university level. *Sociometry*, 1947, *10*, 178-185.

VAN DYNE, VIRGINIA E. Personality traits and friendship formation in adolescent girls. *J. soc. Psychol.*, 1940, *12*, 291-303.

VREELAND, F. M. Social relations in the college fraternity. *Sociometry*, 1942, *5*, 151-162.

WESCHLER, I. R., TANNENBAUM, R., & TALBOT, E. A new management tool: The multi-relational sociometric survey. *Personnel*, 1952, *29*, 85-94.

WHERRY, R. J., & FRYER, D. Buddy ratings: Popularity contest or leadership criteria? *Sociometry*, 1949, *12*, 179-190.

WILLIAMS, S. B., & LEAVITT, H. J. Group opinion as a predictor of military leadership. *J. consult. Psychol.*, 1947, *11*, 283-291.

WINSLOW, C. N. A study of the extent of agreement between friends' opinions and their ability to estimate the opinions of each other. *J. soc. Psychol.*, 1937, *8*, 433-442.

WOLMAN, S. Sociometric planning of a new community. *Sociometry*, 1937, *1*, 220-254.

ZELENY, L. D. Sociometry of morale. *Amer. sociol. Rev.*, 1939, *4*, 799-808.

ZELENY, L. D. Measurement of social status. *Amer. J. Sociol.*, 1940a, *45*, 576-582.

ZELENY, L. D. Sociometry in the college classroom. *Sociometry*, 1940b, *3*, 102-104.

ZELENY, L. D. Status: Its measurement and control in education. *Sociometry*, 1941, *4*, 193-204.

ZELENY, L. D. Selection of the unprejudiced. *Sociometry*, 1947a, *10*, 396-401.

ZELENY, L. D. Selection of compatible flying partners. *Amer. J. Sociol.*, 1947b, *52*, 424-431.

ZELENY, L. D. Adaptation of research findings in social leadership to college classroom procedures. *Sociometry*, 1950, *13*, 314-328.

CHAPTER 12

The Interview: A Tool of Social Science

ELEANOR E. MACCOBY,
Harvard University

AND

NATHAN MACCOBY,
Boston University

INTRODUCTION

The present discussion of the interview will be confined to the uses of interviewing for *research.* The interview has, of course, many other purposes: it is used, for example, for diagnosis and therapy, and for selection of personnel. But it is in the uses of the interview for research that our primary interest as social scientists lies. Using an interview for research purposes implies that the interview is intended to yield some sort of measurement, however rudimentary, along some dimension or variable. Usually, also, the objective will be to relate several variables, thus measured, to one another. In the following discussion then, the interview will be considered from the standpoint of the usefulness of the research data it provides. Some interviewing is, as we shall see later, preliminary to the formal process of measurement, in that it is used to suggest what the relevant dimensions for research shall be.

For our purposes, an *interview* will refer to a face-to-face verbal interchange, in which one person, the interviewer, attempts to elicit information or expressions of opinion or belief from another person or persons. We are excluding both paper-and-pencil questionnaires and behavioral observation from consideration as interviews. It should be noted that behavioral observation may cover observation of either verbal or nonverbal behavior. Not all observations of a subject's verbal behavior may properly be classed as interview data. For example, observations of the verbal behavior of small groups holding discussions, or observations of a parent answering the questions of a child, are not considered interviews within the limits of our definition. In the interview, the verbal expressions of the respondent must be directed toward the interviewer, in response to the interviewer's questions or comments.

In the following discussion of the interview, we shall trace briefly some of the historical origins of present-day interviewing practices, and then consider the place of the interview in research design. Next, we shall discuss some major varieties of the interview (principally standardized versus unstandardized interviews), and then take up the problems and strategy of designing specific interviews, touching upon question wording, and the use of indirect or "projective" questions. Then we turn to the interviewing process itself: the role relationships between interviewer and respondent, the problem of obtaining the respondent's interest and cooperation, techniques of asking questions (including nondirective probing), and finally the process of recording the respondent's replies. Administrative problems in carrying out a field interview study are dealt with briefly. The problems peculiar to interviewing children and interviewing in other societies than one's own are then discussed. The final section of the article is devoted to evaluation of the interview as a research tool: the extent, nature, and control of interviewer error or bias are discussed, along with the problem of validity of interview findings and the merits and demerits of the interview as compared with other methods of measurement available to the social scientist.

HISTORICAL ROOTS OF PRESENT-DAY INTERVIEWING

Historically, the development of the interview as a research tool has two major roots. The use of interviewing in clinical diagnosis and counseling indicated that considerable psychological information could be obtained from interviews. The interviewing techniques which grew up in the clinical setting, however, were not developed with the problems of measurement primarily in mind. Here the emphasis was upon richness of response, and little attention was given to the ordering and classification of the resultant data. An antidote for this absence of concern with measurement was provided by the growth of psychological testing. During World War I, the development of the Army alpha intelligence tests lent strong impetus to the growth of psychological tests and measurements. This period saw the beginning of the use of attitude scales in which the multiple choice test item was adapted to attitude measurement. World War II saw a tremendous increase in the use of this kind of measure. The use of such checklist questionnaires permitted the development of various kinds of scaling techniques that enabled social scientists to obtain objective and reliable measures of attitudes.

In the hands of such research workers, the personal interview was employed largely as a device for administering checklist kinds of questions when it was not possible to assemble respondents for group administration of paper-and-pencil questionnaires, or when it was feared the respondent might not be able to handle written materials. Much public opinion polling and market research interviewing has been essentially an oral adaptation of the objective questionnaire.

In recent years, some psychologists (often those with a Gestalt or Lewinian bent) were interested in exploring cognitive processes through the interview. They felt that the "depth" interview which has been developed in clinical work was more productive for studies of the psychological factors underlying attitudes than the objective type of attitude scale. However, in order to quantify their interview information, they categorized the interview material into codes quite similar to the multiple choices used so widely in attitude measurement. As a result of such attempts, it is now possible to achieve a high level of reliability in the coding of most open-ended or depth interview material (H. Guetzkow, 1950, and Survey Research Center, University of Michigan, 1947).

Interviews as currently employed for research purposes show clearly the influence of both major lines of historical origin. The merging of the test tradition with the clinical tradition has resulted in a wide variety of interview techniques that are now available to the research worker, combining the two approaches in varying degrees.

THE INTERVIEW AS A RESEARCH TOOL

The interview in social psychological research is a general tool which may be employed in a variety of research designs. It may or may not have a place in any particular piece of research. It does not in itself constitute a complete research method. For example, if the object of the research is to test a hypothesis concerning attitude change, a paper-and-pencil test or a set of behavior observations may well be preferable to the use of an interview as the tool for measuring the attitudes of subjects in experimental and control groups. Even if the interview proved to be a desirable method in such an experiment, its application does not in itself constitute the method by which the hypothesis under consideration is tested. It is necessary to formulate a specific design, such as, for example, the "after only" design (Hovland, Lumsdaine, and Sheffield, 1949). In such a design, the interview or the questionnaire or the behavior observation is used to measure the attitudes of the experimental and control groups, and constitutes an important tool in the experiment. Unfortunately, the interview is too often used as a substitute for study design (Stouffer, 1950), in which hypotheses are not formulated nor designs set up which would provide adequate tests of hypotheses.

The interview may have a variety of roles in the design of a study. It may be employed during the early stages of a study to help identify the relevant dimensions, to suggest hypotheses, and to reveal the natural frames of reference existing in the minds of respondents. When paper-and-pencil questionnaires are used as the basic instrument of measurement, the interview may be employed to pretest the questionnaire form.

The second possible role of the interview in the research process is as the main instrument of data collection. Interviews may be used to obtain base-line measures on the experimental and control groups before the application of an experimental treatment, or to measure these groups *after* the application of the experimental procedure, or both. In nonexperimental studies, the interview is commonly used to obtain data for scale measures of a set of variables whose interrelations are to be studied. When the interview occupies the central position as a study's primary instrument of data collection, the problems of standardizing the interviewing technique and measuring reliability and validity become of greater importance than they would be otherwise.

A third role of the interview is to clarify findings which emerge from the use of other techniques. Merton and Kendall (1946) point out that when the experimental variable is a complex one (as it is in much communications research, where the experimental group is exposed to a film or radio program), interviewing may indicate what aspects of the experimental procedure were responsible for observed effects. Furthermore, the interview may offer clues as to why the experimental procedure does not produce the anticipated effects, when unexpected results do occur.

Interviewing often plays an inconspicuous but crucial role in the study of traditional psychological problems. Answers to questions constitute an important part of the data in psycho-physical experiments such as those on brightness cues in depth perception. Interviewing is also an integral part of individual intelligence testing and studies of verbal learning. Sometimes interviewing is employed as a means of inducing stress, which may in turn be measured in terms of various emotional response indicators such as pneumographs, sphygmographs, etc. And sometimes, the interest of the research worker in the interview is not so much *what* the respondent says, as how he says it. During World War II, interviewers for the Office of Strategic Services (OSS Assessment Staff, 1948) were interested in the ease with which a candidate could talk about his early experiences; the data provided by the interview in such a case would not be obtained by the usual content analysis of an interview record, but rather by the interviewer's rating of the respondent's reactions to the interview itself. Similarly, Gerald Caplan, an Israeli psychiatrist, has taken the significant data emerging from an interview to be the role in which the respondent places the interviewer. That is, he observes whether the respondent treats the interviewer as a parent surrogate, as an instrument to be manipulated for his own ends, etc., and is more interested in what such behavior on the part of the respondent reveals about his personality than in the manifest content of the respondent's answers to questions.

It is evident, then, that the interview can occupy many different roles in the research process, and may be used to produce several different kinds of data.

THE STANDARDIZED VERSUS THE UNSTANDARDIZED INTERVIEW

By a standardized interview we mean one in which the questions have been decided upon in advance of the interview, and are asked with the same wording and in the same order for all respondents. The questions used in a standardized interview may be either "open" or "closed." The essential feature of the standardized interview is that the interviewer does not have freedom to word or reword questions, to introduce questions which seem especially applicable to the individual case, or to change the order of topics to conform to the respondent's spontaneous sequence of ideas. In the unstandardized interview, the interviewer's technique is completely flexible, and is varied from one respondent to another.

Both standardized and unstandardized interviews are widely used by social scientists. The distinction is sometimes made between "structured" and "unstructured" interviews (or "active" and "passive" interviews — Kluckhohn, 1945). This dimension is not identical with the standardized-unstandardized dimension, but it is similar, and much of the discussion of structure in interviews applies to our present topic. Unstructured interviewing is, in general, done without a standard list of questions. It has the further characteristic that the questions are as "open" as possible (see the discussion of open questions below), and offer a minimum of guidance to the respondent. Unstandardized interviewing (and/or unstructured interviewing) has been advocated by Rogers (1945), Roethlisberger and Dickson (1939), Piaget (1929), and Kinsey *et al.* (1948), among many others. Merton and Kendall (1946) have advocated a semistandardized interview which they call the "focused interview," and which employs an interview guide with a list of objectives and suggested questions but gives the interviewer considerable latitude within the framework of the interview guide.

The arguments advanced in support of the different techniques may be briefly summarized:

In favor of standardized interviews:
- (a) they incorporate a basic principle of measurement: that of making information comparable from case to case,
- (b) they are more reliable,
- (c) they minimize errors of question wording.

In favor of unstandardized interviews:
- (a) they permit standardization of *mean-*

ings rather than of the more superficial aspects of the stimulus situation,

(b) they are more valid, in that they encourage more true-to-life replies,

(c) they are more flexible.

Let us consider these points in detail. The most fundamental argument supporting the structured interview is a simple one: that one should not use a rubber yardstick. Ultimately, the research worker's objective is to differentiate among his respondents with respect to a set of variables. If he varies his methods of measurement from one case to another, he never knows whether differences obtained among his respondents are due to differences in the measurement technique or differences in the amount respondents possess of the trait being measured. To take an oversimplified example, a research worker who was attempting to measure the amount of color shock individuals display in response to the Rorschach test would not show some of his subjects cards containing a large amount of red and other subjects cards with a small amount of red, for if he did so, he would not be able to attribute differences in response to differences in the personality trait which is presumed to underlie color shock responses — instead, the differences might be due at least in part to variations in the cards to which the individuals were exposed.

Unstandardized interviewing is similar to using a new set of Rorschach cards for every subject. When interviewers are not provided with a standard list of questions, they vary the wording and order of their questions quite considerably, even when they have been well trained concerning the objectives of the study (U. S. Dept. of Agriculture, 1945). And it has been found repeatedly that even what would appear to be minor changes in the wording of questions produce significant changes in the frequency of given responses (Cantril, 1944). The attitudes which a respondent expresses in an interview, or even the "facts" which he reports, will be, to a considerable degree, functions of the particular questions he is asked, and respondent *A* may not reasonably be classified as more internationalistic, or more authoritarian, or wealthier than respondent *B* (on the basis of answers to interview questions) unless both respondents have been asked the same questions.

It should be borne in mind that the research worker is not always interested in comparing one respondent with another along some scale. In exploratory interviewing, he may be interested in finding out what the relevant dimensions are for later systematic measurement, and it is not particularly important that cases should be comparable to one another. At this stage of research, then, the above objection to unstandardized interviewing does not apply.

Some proponents of the unstandardized interview believe in the value of making methods similar from one case to another so that comparison will be possible, but they believe that using standardized wording and question order is not the way to make the measurements comparable. They point out that the same words mean different things to different people, that when one asks a standardized question, one has not standardized the *meaning* of the question to the respondent (Crutchfield and Gordon, 1947). It is better, say the exponents of the "meaning equivalence" position (Kinsey *et al.*, 1948), to use words which have equivalent meanings to different respondents, even if the words are not objectively identical. It is obvious that even the most dogged exponent of objective standardization, when he encounters a respondent who speaks a different language, will not attempt to interview in a language which the respondent does not understand merely for the sake of maintaining formal stimulus equivalence. In practice, the research worker would translate, or get an interpreter. The same problem arises to a lesser degree for different educational levels; the advocate of the unstandardized interview would ask the well-educated respondent about "inflation" and the poorly educated one about "prices going up" on the assumption that these two phrases are more truly comparable than the word "inflation" used for both groups would be. A common solution to the problem of equivalence of meaning in public opinion work (where the emphasis is usually on objectively standardized wording) has been to aim for the lowest common denominator: to assume that the stimulus value of simple words is comparable all along the range of education levels even if this is not true for complex words. But Kinsey found, studying sexual behavior, that there was *no* set of terms which could be understood by all his respondents. Each group had its own jargon, and it was necessary for the interviewer to detect and use these words in order to be understood. Kinsey also felt that it was necessary to vary question order: he wished to follow the widely approved practice of beginning the interview with innocuous items, burying the more embarrassing ones late in the interview, when rapport would be well established. But he found that some items were embarrassing to some groups, other items to other groups, so that he could only standardize the rapport-building qualities of the interview by varying the order of questions according to the respondent's social group.

The extreme form of the stimulus-equivalence position would be that no word means the same thing to two different people, since they will have had different experiences associated with the word (and meanings are the product of all the individual's past experiences with the word), so that in every instance one should attempt to explore the respondent's idiosyncratic meanings, and to find words which are, for him, equivalent to the meanings intended by the definitions of the variables in the research.

The practical difficulties of trying to carry on research in this way are obvious, and no one, so far as we know, has attempted to do it in detail. Furthermore, it is likely that in trying to make equivalent one aspect of the stimulus situation, the research worker introduces an imbalance into another aspect. For example, Kinsey varied question order to obtain equivalence of rapport qualities. However, in so doing, he sacrificed equivalence of the context in which questions appeared, and it may be that some of the group differences which he reports are in some degree due to these contextual differences rather than to differences in actual behavior from class to class.

How can one evaluate the positions discussed above? It is the feeling of the authors that the logic of research requires (at least in the final stages where *measurement* is attempted) that stimuli be standardized in one way or the other: objectively, or via equivalent meanings. Standardizing the objective stimulus should serve very well in most situations; there is great communality of meanings for most words among people speaking the same language. Kinsey's difficulties in finding words with a generally understood meaning stem, we believe, from the special nature of his subject matter. When widespread taboos exist against the written or spoken expression of certain facts and ideas, the conditions are maximal for the development of idiosyncratic meanings and special jargon among the small group who do communicate with each other about the topic. For most research, the problem would be much less acute, and there would be little reason for attempting to go beyond the standardization of the objective stimulus (assuming, of course, that careful pretesting has weeded out most of the words which would have special meanings for any significant proportion of the population being studied).

When the research worker does feel the need to work for special wordings for certain subgroups in the interests of meaning equivalence, it will be necessary for him to keep in mind the following questions:

1. Are the characteristics which place the respondent in a particular meaning category identifiable by the interviewer? If not, the system of equivalents cannot be put into effect. For example, it would be possible to specify that a question should be asked in one form of Negroes and another of whites. But it would be more difficult to provide a special form of a question for people with a father fixation, even if this might be highly relevant to a person's probable interpretation of the meaning of the question, because the members of the group for whom the special wording was to be employed could not readily be specified in advance.
2. Are the equivalent meanings of identifiable subgroups known (or can they be found out as part of the research process)? In the case of translation into another language, or substituting a subgroup's synonym, the problem should not be too difficult, although the meanings will never be exactly equivalent. In other cases, the difficulties are likely to be enormous.
3. Are the gains made by trying to get equivalence of one aspect of the interview large enough to offset the nonequivalences of other aspects which almost inevitably accompany such an attempt?

We have been discussing the implications of standardized versus unstandardized interviewing for producing information which will be comparable from case to case. There are a number of other considerations in the choice between these two methods, for example, their respective reliability and validity. There are two kinds of questions to be asked relative to reliability: (1) when the same individuals are interviewed a second time, which method will yield the closest correspondence between the first and second interview? and (2) assuming that ratings are to be made on the basis of interview write-ups, will the write-up of a standardized interview permit more reliable ratings than the write-up of an unstandardized interview? There are as yet no studies available directly testing these questions. The presumption would be, however, that the standardized interview would be superior on both counts. In the comparison of the first with the second interview, the unstandardized interview would produce differences due to interviewing variability as well as those due to respondent variability, while in the standardized interview error due to variations in interviewing technique should be much less. On the question of reliability of ratings, the interview write-ups from an unstandardized set of interviews usually produce the following kinds of difficulties in the rating process: (1) The interviews tend to be very full in some areas, very thin in others. In fact, without a standardized set of questions before him, the interviewer is very likely to overlook certain topics which he should have covered, and it is usually impossible to rate all the interviews on all the dimensions included in the study. Of course, the use of a detailed "interview guide" such as Merton recommends will help to ensure complete information, even when questions are not standardized. Furthermore, in some situations it is possible to revisit the respondent to obtain items of information which were overlooked in the first interview. When this is possible, the standardized list of questions is less important. (2) Some of the material in some of the interviews will be ambiguous as to whether or not it is relevant to the dimensions of the study. Suppose, for example, that in a study of the public's information about mental health problems one of the topics for discussion in an unstandardized interview is alcoholism. If the respondent talks at length about drunken driving on New Year's Eve by otherwise relatively sober people are his remarks, revealing as they may be of his attitude towards drinking, to be considered as revealing attitudes toward alcoholism? In unstandardized interviewing, the chances are great that the respondents will vary greatly in the frame of reference from which they discuss a question, and raters will vary in whether they consider the frames of reference of several respondents sufficiently comparable to justify rating them along the same dimension. Such variation, of course, produces unreliability of rating.

The relative *validity* of the two methods may be an entirely different matter. A standardized interview is sometimes stilted and unnatural—it may put the respondent on the defensive by making him feel that he is being tested. The unstructured interview, on the other hand, is assumed to be more like a real-life conversation; presumably, it permits the respondent to follow his natural train of thought, and therefore tends to bring out material which is more likely to enable the research worker to predict what the respondent would do or say in real-life situations.

So far, there are no studies on the comparative validity of the two types of interviews. It may be true that some varieties of standardized interviews (particularly the fixed-alternative type) jeopardize rapport. Those who have worked with carefully pretested open-end standardized interviews know that it is possible to maintain the conversational quality of an interview within the framework of a standardized set of questions, at least to a considerable degree, but the verbal interchange between interviewer and respondent is never as free as it would be without a standard list of questions. It remains to be seen whether unstandardized interviews have sufficiently greater validity so that ratings based upon them will predict criterion variables better than will ratings based on standardized interviews.

A point often made in support of unstandardized interviews is that they are flexible. By flexibility is meant that the interviewer may adjust his technique to the individual respondent's particular situation. He may explore unusual ideas which the respondent mentions, omit topics which do not seem applicable in a particular case, take advantage of extraneous events which occur during the interview to introduce a related topic, etc. It is true that it is very difficult to design a standardized interview in which all questions are applicable to all respondents, but a certain amount of flexibility may be achieved within the framework of a standardized interview by providing subsets of questions which are to be asked only of particular groups of respondents—for example, only of those who have had direct contact with members of a minority group, or only those who have certain minimum information about the U. N., etc.

So far, in the discussion of standardized versus unstandardized interviews, the two techniques have been presented as though they were mutually exclusive alternatives, with no middle ground. Actually, for some purposes, a semistandardized interview may be best. For example, in the interests of flexibility, the research worker sometimes specifies exactly a series of main questions, which are to be asked of everyone, but lists a series of optional subquestions or probes which the interviewer is free to use or omit, depending upon the respondent's answers to the main questions.

The research worker's choice of a fully standardized, semistandardized, or unstandardized interview will depend very much on the stage of development of the subject-matter area in which he is working. It is fruitless to attempt standardized interviewing before the dimensions of an area are well delineated, and a good many years of exploratory work are often needed before the precision of standardized interviewing is possible or justified. (The insights of Freud, arising out of his completely unstandardized interviewing, are a lesson to the overzealous researcher who would begin late-stage research too soon.) Nevertheless, the obligations of measurement cannot be evaded indefinitely by labeling research "exploratory." Every developing field in social science must arrive at the measurement stage eventually, at which point the standardized interview comes into its own.

The Phrasing of Interview Questions

The problem of selecting questions and wording them properly is essentially the problem of translating the objectives of the study into the concrete set of stimuli which will determine the nature of the final data. It is precisely at the point of wording questions that objectives may be garbled or lost sight of completely.

A temptation which besets research workers is to translate objectives into questions too directly. If an objective of the research is to find out something about the determinants of a given attitude or piece of behavior, the research worker may simply ask the respondent "Why did you do that?" or "Why do you feel that way?" The weaknesses of the simple "why" question have been ably described by Lazarsfeld (1935), who pointed out what a wide variety of frames of references it might elicit, only some of which are relevant to a study's objectives. In many cases, its worst weakness is that it asks the respondent to be the scientist, asks him to provide insights about the determinants of his own behavior which he should not be expected to provide. The research worker may want to know "why" certain behavior occurs, but he can seldom find the answers simply by asking respondents to explain themselves. The very question "why?" implies an explanatory system —a set of categories which will be related to the phenomenon under study. For example, suppose the research worker wanted to find out why the members of one group display more overt aggression than the members of another group. What are some of the possible levels of explanation? The group may be aggressive:

(1) because they hate and fear others,
(2) because they are frustrated,
(3) because they believe others hate and fear them, and therefore feel justified in attacking first,
(4) because they have no insight into their own motives,
(5) because they were severely punished in childhood and have "identified with the aggressor,"
(6) because they are larger and stronger and have learned that aggression will get them what they want,
(7) because they have an internally unbalanced social system and can only keep it intact by creating an "out-group" to attack.

Clearly, there are different levels of explanation here. Reason 2 may lie behind reason 1. If we want to explain aggression in terms of projection, we might expect that the following sequence is involved: the individual is frustrated, he is prevented from showing aggression toward the instigators of his frustration, and then, provided he has no insight into his own motives, he projects (attributes his own aggressive feelings to others), and then attacks them in "self-defense." From the point of view of this presumed sequence of events, (2), (5), (4), and (3) are all parts of the explanatory sequence. None of them is *the* explanation. Reason 7 comes from a completely different explanatory system—a system of concepts arising out of the theory of group process.

The research worker must choose the system of explanatory concepts which he wishes to employ, and then must ask questions in his interview which are relevant to these concepts. If he chooses one set of concepts, he will ask about the severity of discipline meted out by the respondent's parents. If he chooses another set of concepts, he will ask about the relations of group members to one another and to their leader. He will seldom succeed in getting at desired levels of explanation merely by asking the respondent why he does what he does or believes what he believes.

The research worker, then, decides what questions to ask by using all his ingenuity in selecting kinds of explanations which offer promise of predicting the behavior he is studying, and then finding specific questions which will help to measure the dimensions he has chosen. Even after he has decided that he must measure "severity of current frustration," "severity of punishment in childhood," "cohesiveness of the group," or "degree of insight," he has the very great problem of translating these dimensions into questions which will provide reliable and valid measures. An interesting example of translating an underlying dimension into specific questions has been provided by the students of authoritarian personality (Adorno, *et al.,* 1950), who ask about attitudes toward women in politics, and the respondents' views about the proper reaction when one's honor is insulted, etc.—questions which a researcher would ordinarily not think of asking in a single study. They appear together because the research worker has designed questions which are all related to an underlying personality dimension which unites the seemingly unrelated bits of behavior.

Whether the research worker devises a set of standardized questions in advance or composes his questions as he goes along in the interview, he must be cognizant of the pitfalls in phrasing questions. Stanley Payne (1951), in his book, *The Art of Asking Questions,* has discussed many of the problems of wording, and has provided a checklist of one hundred considerations to keep in mind when preparing questions for public opinion interviewing. Other discussions of question wording which present both summaries of experimental findings on the effects of variation in question wording and some of the "wisdom" of practitioners in the field may be found in the literature (Cantril, 1944; Parten, 1950; Jahoda, Deutsch, and Cook, 1951). It is not possible here to present a complete and specific list of the do's and don't's of question wording. Instead, we wish to discuss certain major problems which the research worker faces in phrasing questions, and to summarize some of the solutions which have been found workable.

A primary problem in wording questions is to make sure that the questions create for the respondent as clear a picture as possible of the subject matter under discussion. Of course, in some instances, a truly clear picture is impossible, because the respondent is completely uninformed about the topic and completely uninterested in it, so the question can hardly provide a frame of reference which will be meaningful for him. Frequently, the research worker will employ an information filter question, which will permit him to identify such people and omit further questions for them.

The advantages of having questions worded as clearly as possible are perhaps too obvious to need recounting. In the first place, when questions are ambiguous, standardization of the stimulus has not been achieved, because some respondents are talking about one thing and others about another (Crutchfield and Gordon, 1947); in the second place, interpretation and classification of the answers is hampered, because there will be too many answers which do not fit the intended scales, and too many "don't know" answers.

Following are some of the rules of wording which have been worked out in the interests of question clarity:

(a) Avoid words with a double meaning. The word *fair,* for example, may mean either "equitable" or "not very good." The phrase *put up* may mean to nominate a candidate, stay for the night, endure an insult, or preserve fruits.

(b) Avoid long questions. The respondent will not be able to keep the entire question in mind, and will answer only part of it, ignoring some of the qualifications the research worker intended.

(c) Specify exactly the time, place, and context which you want the respondent to assume. A question about "When you were in your teens" may elicit very different answers from "When you were young" or "when you were still in school."

(d) Either make explicit all the alternatives which the respondent should have in mind when he answers, or make none of them explicit. For example, if you ask: "Are promotions in your office usually based on merit?," the respondent many answer "yes" hastily without considering that raises might be based on seniority, favoritism, etc. If the interviewer reminds the respondent of the other possibilities, a more thoughtful reply will be forthcoming. For some purposes, it would be better to mention none of the alternatives. The question could then be phrased: "What are promotions in your office usually based on?"

(e) When the interview concerns a subject with which the respondent may not be very familiar, or one in which he may not have the necessary technical vocabulary, it is sometimes desirable to preface questions with an explanatory paragraph or an illustration which will set the stage for the question the interviewer wants to ask. (See NORC, 1950.) The question itself may then still be kept brief.

(f) It is often helpful to ask questions in terms of the respondent's own immediate (and recent) experience, rather than in terms of generalities. For example, in studying supervisory practices, the interviewer may ask, "Think back to the last time one of your workers came in late. What did you say?," rather than, "What do you usually say when an employee comes in late?" Similarly, studies on exposure to mass media have found that when people are asked how much time they spend listening to the radio or looking at television *on the average,* they are likely to underestimate the time. More accurate estimates can be obtained by asking people to say how many hours they spent in this way on the preceding day, or what specific programs they saw or heard during a given period of time.

When an interview deals with sensitive topics, attention must be given to wording questions in such a way as to minimize ego defenses. As will be seen later, research workers sometimes have recourse to projective or other kinds of indirect questions. There are ways, however, in which direct questions may be worded so that the respondent will feel relatively free to report candidly. Essentially, the problem is to keep the respondent from feeling that certain answers will mean loss of prestige. This can be accomplished by making the answers which the respondent might want to give seem socially acceptable. The following are some ways in which this has been done:

(a) Indicate that other people hold what might otherwise be considered a frowned-upon opinion. "Some parents feel that it's terribly important to teach a child not to fight with other children. Others feel that in some circumstances, a child must learn to fight. How do you feel about this?"

(b) Balance against each other alternatives each having a positive value loading: "Do you and your family believe in showing your affection for each other a lot, or are you fairly reserved sort of people?" An earlier version of the question, couched solely in terms of whether the respondent's family was affectionate, yielded a bunching of cases at the "very affectionate" end of the scale.

(c) Assume that the respondent possesses the low-valued attitude or behavior in some degree, and place the burden of denial on him. Kinsey, *et al.* (1948) use this technique when they phrase their questions, "When did you first . . ." instead of "did you ever . . ." Similarly, if people are asked whether they paid cash for an article or bought it on the installment plan, too many say they bought for cash. But if the question is worded, "How much are your monthly payments?," the proportion who say they paid cash corresponds much more closely to the proportion known to be true from sales figures.

(d) Use euphemisms. For example, in an interview on child-rearing, "training methods" or "methods of discipline" would be better than "methods of punishment."

(e) If the sensitive area involves the expression of criticism of a person or institution, provide an opportunity for the respondent to voice praise first, so that he will not feel he is being unfair. "What do you like best about your supervisor?," then "What do you like least?"

(f) Introduce face-saving phrases and words. For example, "Do you happen to know who the Congressman from this district is?" was used instead of the more abrupt "Do you know . . .?," and Russians escaping to the West, who might hesitate to admit former Communist Party membership, were asked: "Were you able to keep out of the Communist Party?"

One of the major areas of discussion among social scientists who have used the interview as a research tool centers around the problem of whether questions should be "open" or "closed." An open question gives the respondent little or no guidance as to the form or content of his answer. A closed question, on the other hand, limits his answer severely: it specifies a series of alternatives (sometimes only two) from which he is to choose. The large polling agencies commonly use closed questions. For example, Roper frequently employs the multiple-choice (or "cafeteria") question, in which the respondent is given a card containing a series of alternatives, and is asked: "Which of these comes closest to your own opinion?" The Gallup questions on voting intentions, and most of his other questions, are closed, and are frequently of the two-choice type; the respondent is to say "yes" or "no," agree or disagree, approve or disapprove, or choose one of two candidates. Closed questions are commonly used in studying consumer purchases. The consumer is asked which of a list of household durables he bought within the past year; he looks over the list and designates the items which he bought.

Open questions state a problem without specifying a list of alternative answers. Following are some examples of open questions: "How would you compare the way you are raising your children with the way your parents raised you?," "It has been suggested that the United Nations should have the power to make laws which the people in the United States and the other member nations would have to obey. How do you feel about this?," "Did you buy any large items for your home during the past six months?" (If yes) "What things were those?"

It should be borne in mind that the openness of a question is a matter of degree. Questions vary from those which restrict the respondent's response completely to those which do so minimally. Every question limits the range of possible answers to some extent, by virtue of specifying the subject matter under discussion.

For example, in interviewing about prices of consumer goods, the following question might be used: "Would you give me your comments about the prices of the things you buy?" This question is very open, and permits a range of answers which would include, "My grocer will overcharge if I don't watch closely," "Eggs have just come down, but butter is still high," "Things were much cheaper when I was a child," and "I think they'll be coming down soon." If the research worker is interested in price *expectations,* he may use a question which is less open but still not closed: "What do you think will happen to the prices of the things you buy during the next year or so?" The completely closed version would be: "What do you think will happen to the prices of the things you buy: Will they go up, down, or stay about the same?" Here, the respondent is almost fully restricted as to the nature of his answer.

It will be seen that open questions, in a sense, ask the respondent to *recall* something — to produce it spontaneously. Closed questions, on the other hand, ask the respondent to *recognize* something. The literature on recall versus recognition memory tells us that more will be recognized than will be recalled, and such proves to be the case with open versus closed questions. In consumer purchase studies, for example, the open question yields fewer buyers of any given item than does the closed question. The chances are that the closed question is more *valid* in purchase studies; that is, that the housewife who looks down the list and says, "Oh yes, I did buy an electric iron. I had forgotten about that," really did buy the iron. On opinion questions, however, the difference in validity is by no means so clear. The open question may produce errors of omission, but the closed question may produce equally great errors of commission. If the respondent is presented with a series of alternatives, he may choose one even though he has never thought about the issue before, and has no opinion or information on it. Choosing something saves him from having to admit his ignorance. In the open question, where he must produce an answer in his own words, the chances are probably better that ignorance will be revealed and that the spontaneous answers which do emerge will reflect his more considered opinions (Campbell, 1946).

An advantage of the open question, then, is that it permits the research worker to detect ignorance, misunderstandings of the wording of the question, unexpected frames of reference, etc. Furthermore, it is not so likely to suggest an answer to the respondent. The open question is presumed also to have the advantage (although this, again, has not been documented) that it promotes rapport because the interchange between interviewer and respondent seems more like an ordinary conversation. With closed questions, the respondent may become irritated by having the words of the fixed alternatives put into his mouth. He may feel that *none* of the answers provided properly represent his opinions, with the nuances he would like to include, and he may prefer to state his answers in his own words. At any rate, open questions are clearly a more natural form of verbal communication and, as such, should produce a closer replica of the kinds of statements the respondent might make in real life.

The primary advantage of the closed question is that it produces greater uniformity among respondents in the dimension along which their answers lie. In the series of questions about consumer prices, given earlier as an example, it will be evident that the first form of the question is much too open if the objective of the study is to learn something about price expectations. While some of the respondents happen to answer the question from this frame of reference, many do not, and their answers would not be useful in analysis of the data. The closed question pins down each respondent and obtains an answer which fits into the categories of analysis. It may be objected that to force individual answers into predetermined categories will not serve the purposes of analysis if this has been done at the expense of validity — that is, if it has been done by forcing an opinion on the respondent which is not really his. This is true, of course, and careful pretesting of fixed-alternative questions must be done to make sure that the question itself is meaningful to most respondents and that the answers provided offer a reasonable range of answers so that the respondent can find his own position represented. The reason that many answers to open questions are not useful in analysis is usually not that the respondent could not have answered along the relevant dimension; usually, he can and will give an answer which fits the categories intended for analysis if it is made clear what the desired dimension is. For example, in the case of price expectations, the woman who answered the original question in terms of her memory of prices in her younger days could no doubt talk about the probable trend of prices in the future — she only happened to take the original question in another light. Those who have used open questions have found that in many instances they must follow them by more specific, closed questions, in order to make sure that all respondents give material relevant to the variables which are being measured in the study.

Another advantage of the closed question is that it facilitates field coding. When the question limits the possible answers to a specified set of alternatives, the interviewer may simply check the alternative which the respondent chooses and, in many instances, the answer may be transferred directly from the interview form to a punch card. When open questions are used, on the other hand, and the respondent answers at length in his own words, there are considerable problems of classifying and using the answer in analysis. The interviewer may be instructed to classify the answer into a set of categories as soon as the respondent has finished giving it. For example, in reply to a question on sovereignty for the United Nations the respondent may say: "Well, I don't think anybody should have the right to tell the United States what to do. After all, we're footing the bill for all those other countries. Of course, I believe in the United Nations — do they have to be able to make laws to keep going? If they do, I suppose I'm in favor of it." If the interviewer is provided with a set of categories: "Approve, approve with reservations, pro-con, disapprove with reservations, disapprove" and is required to check one, it is evident that he will have difficulty, and that the reliability of his judgment will be questionable. On the other hand, if he is *not* provided with categories, and is instructed to write down the respondent's answer verbatim, it will be necessary to classify those answers in a later step of analysis. Coding free-answer material is costly and time consuming.

Practitioners in the field of public opinion interviewing nearly all use some open and some closed questions. It is widely recognized that each technique has its own virtues. The choice between forms depends on the objectives of the study, the time and money available, and perhaps most importantly, on the level of the research worker's knowledge about the dimensions of opinion in the area he is studying. If he is only beginning to explore an area, he must use open questions, for he will not know how to phrase alternatives for closed questions. If he knows the specific manifestations of the dimensions he is studying, and is satisfied to obtain data on these dimensions and nothing else, then the fixed-alternative question may serve his purposes very well.

Many studies include a range of questions, some open, some closed. It has been common to arrange groups of questions in a "funnel" — that is, to begin questioning on a particular topic with an open question, and then follow with closed questions on specific points of interest to the research worker. The primary advantage of this technique is that it permits study of the spontaneous frame of reference which a respondent will adopt before any line of thought has been suggested to him, while at the same time providing comparable data for all respondents on specific points. In a recent paper, Gross and Mason (1953) have suggested that there are instances in which the research worker should use the opposite technique: the "reverse funnel." Gross and Mason were studying role conflict in school superintendents, and embarked on their work using open questions. Using all their ingenuity in phrasing questions, they could not get responses related to role conflict with open questions during their pretest. Instead, the questions brought out comments about situations in which the superintendent was in disagreement with groups or individuals in his community. After considerable experimentation, Gross and Mason found that they had to reverse their procedure. They first presented the superintendent with four structured situations which all superintendents face (salary increases for teachers, etc.). Then a series of fixed alternatives were given, concerning how various people or groups might expect him (the superintendent) to act. Then the superintendent was asked whether the individuals who expected different things of him had a *right* to do so. If different expectations were perceived as justified, then the case was considered one of role conflict. After a considerable period of highly structured interviewing in the manner described above, Gross and Mason found that they could ask open questions about *other* role conflict situations which the superintendent had encountered, and get meaningful answers.

Gross and Mason's experience seems to suggest that there are instances in which the research worker wishes to inquire about a general category of experience, but the category is one which is not usually seen as a category by the respondent. However, after giving a series of instances of the category he has in mind, the interviewer can impart to the respondent a sense of the meaning of the category (assuming that the instances actually do have something in common!), and permit him to talk about other instances and finally come to a generalized opinion.

INDIRECT VERSUS DIRECT METHODS OF QUESTIONING

We have discussed elsewhere the fact that respondents are sometimes unwilling or unable to report directly the data a research worker wishes to secure. He may wish to measure certain underlying personality traits (such as "authoritarian personality") which the respondent could not describe in himself even with the best possible rapport between interviewer and respondent. Or he may be interested in measures of "sensitive" areas, such as sexual behavior, financial situation, intergroup prejudice, or neurotic symptoms. Presumably, the better the rapport the interviewer is able to achieve with the respondent, the more valid will be the answers he will get in sensitive areas, but usually there will be some distortion of the individual's report in the direction of what he believes to be socially acceptable.

A variety of indirect techniques have been used in interviewing when the subject matter of the study is likely to arouse resistance. The rationale behind the indirect techniques is that they will produce relevant information without the respondent's being fully aware of the implications of his replies.

In recent years, there has been great interest in "projective" questions as devices for indirect measurement. While the term is variously used, projective questions usually refer to questions asked about someone other than the respondent. The assumption is that the respondent will put himself in the place of the other person referred to in the question, so that his answer will reflect his own attitude. Thus, if a respondent says that "other people" probably don't believe in God as strongly as they say they do, the answer would be taken to mean that the respondent's own private belief in God is weaker than his public belief. The principle of the projective measure has been made familiar through the widespread use of such clinical instruments as the TAT and Rorschach test

Projective questions have probably been used less widely in research-oriented interviews than they have in paper-and-pencil personality inventories and in diagnostic and therapeutic interviews. In public opinion work, projective questions have given rise to considerable difficulties in interpretation of the answers. In some instances, answers do seem to reflect the respondent's own attitudes, but in others they seem to represent a very different thing: his realistic appraisal of someone else's attitudes. For example, a young woman being interviewed about her attitudes toward her job situation was asked: "How do most of the girls in the office feel about the supervisor?" She answered: "They think he's wonderful. They'll do anything for him." At which point the interviewer followed with: "And how about you — how do you feel about him?" with the reply: "I really detest him. I'm trying to transfer out of the unit."

The problem of validating projective questions is essentially difficult. One can, of course, correlate the answers to projective questions with the answers to direct questions on the same subject. But one should not expect a very high correlation — research workers have had recourse to projective questions in the first place because they hoped these questions would measure something which direct questions cannot measure. The hope has been, in other words, that indirect questions would yield information that was more *valid* than the information produced by direct questions. The literature is almost devoid of evidence that this hope is at all justified (Campbell, 1950, p. 30).

One study (Getzels, 1950) provides some support for the greater validity of indirect questions, although it was done with a paper-and-pencil questionnaire and used sentence-completion items, and may therefore not be entirely applicable to the face-to-face interview. Getzels compared direct items, such as "When I am asked to be in charge, I generally tend to . . .," with indirect (or projective) items such as "When they asked him to be in charge, he . . ." He found that indirect items on neurotic symptoms would discriminate between a group previously classified as "maladjusted" and one classified as "adjusted" *better* than would direct items. He also found that the difference between the answers to direct and indirect items depended upon how much the topic was subject to clearly perceived public norms and social pressures. His indirect items yielded more expressions of hostility toward Negroes, more expressions of conflict with parents, more "neurotic" answers to personality-relevant questions, than did direct items. On the other hand, there was little or no difference between the responses to direct and indirect items in political-economic questions where a variety of answers would be equally acceptable. A further finding was that while the answers to direct items in areas subject to social norms were grouped around the conforming end of the scale (in a *J*-curve), the answers to the projective questions in the same areas were more normally distributed.

Getzels asked his subjects about their perceptions of the person referred to in the indirect questions. Usually, they reported that they thought of the person as "someone like myself." The attitudes which they attributed to this mythical other person were very different from the attitudes which they attributed to "a cross section of the people." Getzels has some evidence, then, that his respondents were identifying with the subject of the indirect item, and not merely describing what they believed other people's opinions to be. His findings on the validity of indirect items apply to questions which refer to indefinite subjects ("he," "John") and *not* to questions referring to "most people" — the form more commonly used as projectives in public opinion work.

In personal interviewing, it is difficult to ask projective questions such as "When someone is rude to John, how do you think he usually feels?" The respondent may ask who John is. On the other hand, if the interviewer asks projective questions in terms of what "most other people" think, the respondent may give his appraisal of the beliefs of others (or say he doesn't know what other people think about the issue), instead of identifying with the subject of the question and projecting his own beliefs in his answer. Getzels' work shows, in fact, that people very often do not project when the "other people" form of a question is asked.

One possible solution in personal interviewing is to phrase questions in terms of "a person": "How do you think a young person would feel about it if his father wouldn't let him stay out on Saturday evenings past twelve o'clock?" Or it is possible to invent a story in which the situation of a mythical person is described, after which the respondent may be asked how he thinks the subject of the story would feel and behave.

An interesting example of the successful use of projective questions in interviewing a sizeable sample is provided by Sanford and Rosenstock (1952). Their objective was to obtain data relevant to authoritarian personality structure. They used a series of cartoon-like drawings, depicting situations where issues of leadership or authority were involved. For example, in one picture a man is standing in front of a group and saying: "Since I'm head of this group, you'd better do as I say." One member of the group is shown answering, but his answer is left blank. The interviewer shows the picture to the respondent, and says: "Here is a picture. One person is talking, the other is going to answer. What do you think the answer would be?" or "Something has been said. What's the first answer that comes to your mind?" The pictures were mixed with a series of more conventional interview questions, including some from the *F*-scale (on authoritarian personality). The entire interview took over an hour.

Sanford and Rosenstock report that respondents were very interested and cooperative in the use of the pictures; they believe that the change from conventional questions served to break the monotony of the interview. It was possible to score the picture responses (using a system of coding responses into categories) with reliability equal to or greater than that for conventional questions, and the reliability from repeated interviews also compared well. Although it was not possible to obtain behavioral data against which to validate the responses to the pictures, the picture responses did discriminate between groups rated authoritarian or equalitarian on the *F*-scale. The technique is therefore promising from the standpoint of validity, and offers the great advantage that it can be employed in interviewing *all* levels of the general public. Burleigh Gardiner (1952) has used a similar technique with comparable success in commercial research.

Projective questions are not the only kinds of indirect questions which have been employed in interviewing. Campbell (1950) and Weschler and Bernberg (1950) have summarized much of the literature. We can only mention a few of the indirect methods here, to illustrate the variety that is possible:

(1) The error-choice method (Hammond, 1948). People are asked to choose among various possible answers to a factual question (such as what proportion of the consumer's dollar goes to profit). All the answers provided are incorrect. The *direction* of the error a respondent makes will reflect an attitude.

(2) Information tests. It has been widely found that the amount and kind of information an individual has about a subject will be a function of his attitudes in the subject-matter area (Newcomb, 1946). Information questions in an interview may be used, therefore, as an indirect measure of attitudes.

(3) Argument completion (Murray and Morgan, 1945). The respondent is given a card containing a description of two people having an argument. He is asked to complete the argument. A series of cards are presented as a test of the respondent's powers of argumentation.

(4) Word association. Vicary (1948) has used the word-association device with some success in market research. For example, the respondent may be presented with a stimulus

word, such as a kind of commodity, and asked to respond as quickly as possible with the brand name that comes to mind.

We have pointed out some of the kinds of indirect questions which have been used in interviewing, and have presented fragmentary evidence that such questions (1) yield more differences from direct questions the greater the social pressure against direct expression of the attitudes, and (2) are more *valid* measures than direct questions when social pressures are present. Balanced against these tentative findings is the fact reported by Campbell (1950) that direct tests tend to have higher reliability coefficients than do indirect ones. A reasonable conclusion appears to be that indirect questions should be used only when the subject matter is one involving strong social pressures against the expression of certain attitudes, and when measures can be taken to strengthen the reliability of the indirect questions sufficiently so that their potentially greater validity will have an opportunity to be effective.

ROLE RELATIONSHIPS IN THE INTERVIEW

One could imagine a measurement situation in which a respondent sits alone in a room where questions are emitted by a phonograph and the respondent's answers are recorded by a mechanical recorder. An interview in which the interviewer is present to ask questions is very different from the hypothetical situation just described, and the difference lies mainly in the nature of the interpersonal relations between the two people involved. The behavior of each member of the pair is determined to a considerable extent by his perception of the other: his perception of the other's status and power position, the other's probable views, and the other's similarity to people with whom he interacts or would like to interact.

What are some of the roles in which respondents may perceive an interviewer? Much depends, of course, on the auspices of the study and the setting of the interview. If the study has been sponsored by a prestigeful institution and covers topics on which the interviewer might be assumed to have expert knowledge, the interviewer may find himself placed in a role similar to that of the family doctor: he is consulted for his advice on the respondent's problems. Such a role for the interviewer may interfere seriously with his purpose of gathering certain kinds of information from the respondent. When the study is sponsored by an industrial concern and the respondent is an employee, there is danger that the respondent will perceive the interviewer in the role of "company man" or informer, and in these circumstances the respondent will not, of course, be candid. When the study has no particular sponsorship, and is introduced at the respondent's door by an interviewer who says he is "doing a survey," the respondent may believe he is a salesman, a solicitor, or an investigator with sinister purposes. Almost always, the role in which the respondent places the interviewer will have some status implications: that is, it will be a role which has more prestige, the same prestige, or less prestige than the respondent's own role, and the nature of the communication which can occur between the respondent and the interviewer will be greatly influenced by the mutual perceptions of status. In general, we know that people are more anxious to communicate to those above them in the hierarchy than to those below them. Back, *et al.* (1950) and Riesman and Glaser (1948) have pointed out that this applies to the interviewing situation. While upper-class respondents feel they have little to gain by expressing their opinions to an interviewer, lower-class respondents are pleased to be consulted. The *content* of the communication, of course, will be affected by the status relationships: the person of lower status will be motivated to present himself in a favorable light to someone who might be in a position to influence his future. At the same time, there may be elements of hostility based on the resentment of authority.

Studies on interviewer bias (to be reviewed later in this chapter) have dealt primarily with the effects upon interview data of the characteristics of the interviewer *independent* of the characteristics of the respondent; the above discussion of role suggests that it would be fruitful to analyze interview data in the light of interviewer and respondent characteristics taken jointly. That is, a young interviewer should exercise a different kind of influence upon the interview, depending on whether his respondent is young or old. Similarly, the effects of the social status of the interviewer, or sex of the interviewer, can be understood only if the interviewer's status or sex is seen in relation to that of his respondent: certain kinds of communication should occur most easily if the respondent perceives himself to be higher than the interviewer on the social scale, certain kinds if lower, and other kinds if the two are at approximately the same level.

The essential point regarding role relationship in an interview is that the interviewer must occupy *some* role, whether he wishes to or not, and that therefore the research worker must be conscious of the various roles possible for his interviewers and attempt to establish the role which will best further the purposes of his study. Having decided upon the most advantageous role, he can govern his choice of interviewers, of sponsorship, and of introductory remarks accordingly. Of course, it is not always possible to anticipate the respondent's role perception of the interviewer, but much can be done to encourage desired perceptions and avoid undesired ones.

What is the ideal role for interviewers? There is no general answer to this question. There are some studies where the role of "expert" may be best, and others where the role of naïve learner may produce more valuable information. There do seem to be some general points regarding the optimum interviewer role which emerge from a variety of studies, however. First, the interviewer should not appear to be completely naïve on the subject matter of the study, or the respondent will feel that his answers cannot be understood. The more specialized the subject matter, the more must interviewers be selected or trained to possess a minimum of sophistication about the topic. For example, Gross and Mason (1953) found that in interviewing school superintendents, interviewers were unsuccessful unless they had an institutional sponsorship which would imply their having some knowledge of educational administration; Kinsey, *et al.* (1948) found that interviewers working with respondents from the underworld must know some underworld jargon to gain the confidence of the respondents, and the Department of Agriculture found that interviewers (especially if they were men) could achieve better rapport with farmers if they had some agricultural experience. On the other hand, the interviewer must not appear to know everything that the respondent knows: he must appear to need, and want, information which the respondent is in a better position to have. Ideally, then the interviewer should appear to be (to some degree) an "expert" seeking additional information from an informed person who has detailed knowledge or opinions to contribute (Bauer, 1953).

An essential element of the interviewer's role is that he should not be in a position to control sanctions affecting the respondent. He should be *outside* the power hierarchy in which the respondent normally finds himself. If he is not, the respondent will attempt to put himself in such a light as to gain the interviewer's support and avoid his censure, and will not talk freely about any topic where he fears he might discredit himself. Furthermore, any in-group applies sanctions against the expression of aggression within the in-group, and if the research worker wants to detect elements of internal hostility and tension, he will be impeded if he is identified as a member of the in-group. It is common knowledge among research workers using the interview that under proper conditions people will talk to an interviewer who is a complete stranger more freely than they would to a personal friend or fellow-worker (England, 1949). The interviewer's detachment from the respondent's real-life situation is an important asset, and in most studies his role must be played in such a way as to maintain this detachment. Miller (1952) has warned about the dangers of "over-rapport." He found that when he established close personal friendships with informants, it became very difficult to question their basic attitudes. A further difficulty is that if the research worker becomes intimate with one person or group among his potential respondents, other respondents will identify him with that person or group and govern their communications accordingly. All this means that, in general, social contacts between interviewers and respondents are not to be encouraged, unless the interviewer is working at the exploratory level and must rely upon participant observation to check some of his interview findings; even then, his research may be hampered by his loss of the stranger role.

The "deeper" the material touched upon in an interview, the more familiar the interviewer should be with the emotional involvements that may appear with communication of such material. Therapists have long been familiar with the strong positive or negative feelings toward the interviewer which may occur. In most social science research work, interviewing is seldom so deep and so prolonged that the interviewer has serious problems of resolving transferences, but he must be prepared for less profound varieties of interpersonal stress. The respondent may exhibit dependent responses toward the interviewer: ask for help, advice, praise. The interviewer must attempt to deflect these demands upon himself without appearing to reject the respondent. If, under the pressure of anxiety aroused by the interview, the respondent becomes hostile, the interviewer must refrain from becoming angry or hurt in his turn, while not appearing indifferent.

It would seem evident that role theory and research can contribute considerably to the understanding of role relationships in the interview, and the interview situation provides a good laboratory for the study of role problems.

GETTING THE RESPONDENT'S COOPERATION

One of the problems in conducting most interview studies is that of getting the respondent's cooperation in being interviewed. While the best methods for getting rapport vary with the time, place, subject, and identity of the respondent, there are a number of general points that can be made that apply to most interview situations. If, for example, interviewing is to be carried out in a community, a carefully planned program of obtaining cooperation within the community is desirable. Depending on the subject matter, various local groups and relevant officials may be interested in the study and can help obtain community acceptance. It should be pointed out, however, that it is often dangerous for the researchers to be too strongly identified with one group in the community. If, for example, an industrial study is being carried out and workers are to be interviewed, support for the study should be obtained from union as well as company officials. Often the fact that such a survey is being conducted is sufficiently newsworthy to attract space in newspapers or other local publications. Even if these media do not directly reach all respondents, an interviewer can be armed with a copy of the newspaper story, and this will sometimes simplify his entrée considerably. Some research organizations find it desirable to make it a rule to notify local police in order to avoid potential difficulty.

One method of attempting to reduce the refusal rate and ease the interviewer's problems in obtaining cooperation from respondents is the use of preceding letters or telephone calls. These devices have sometimes been found to be helpful in the case of respondents in the upper socioeconomic levels. Similarly, when the study is being conducted among the members of an organization (business, government, student body, etc.), the making of appointments with respondents in advance of interviewing can facilitate matters greatly.

Appointments have the advantage of permitting the respondent to select a time for the interview when he will be free and uninterrupted. They also enable him to consider in advance the subject matter of the interview, so that his answers to the interviewer will not be so much "off the top of his head." There may be some disadvantages to making appointments in advance, however: in order to telephone or write, the survey organization normally has to identify the potential respondents, and it is difficult in these circumstances to maintain an atmosphere of anonymity. Furthermore, if telephone contacts are made, it is easier for the respondent to refuse to be interviewed. (In an unpublished study, Peter Rossi found a higher refusal rate with telephone than face-to-face contact.) A compromise solution adopted by some organizations (e.g., Survey Research Center) is to write a letter to the "Head of the household" at a sample address, introducing the study and the interviewer, and giving the time when the interviewer will make his first call, but not giving the respondent an opportunity to refuse until the face-to-face contact has been made. In some studies, of course, it is better not to forewarn the respondent about the interview. In expectation of being interviewed, he might inform himself about the topic, and thus measurements of information level would be misleading.

The above are suggestions for improving public relations in undertaking a survey in a community. In addition, there are a number of things an interviewer can do to gain a respondent's cooperation in the face-to-face situation. First of all, most people like to be interviewed and will usually agree without much persuasion. If the study is serving an important public or scientific need, the appeal to the respondent's altruism is often effective. In any case, the respondent, if he wishes, is certainly entitled to know, whenever possible, the nature and purpose of the study on which he is being interviewed. It should be pointed out, however, that one of the most common errors interviewers make is to give too long and too complex an introduction. Most respondents are not interested in getting complete statements on the nature, purpose and design of the study, and brief introductions work best with such respondents.

Interviewers who are confident and expect to get respondent cooperation usually have the fewest refusals. When interviewers experience several successive refusals, it is wise to have them obtain help on their approach before proceeding to further interviews. Although all interviewers, if they do enough interviewing, will occasionally run into an uncooperative respondent, survey organizations usually find that a large proportion of refusals and semirefusals come from a disproportionately small number of interviewers.

There is considerable evidence to suggest that careful selection and training of interviewers can keep the refusal rate low. When refusals reach fairly high rates of 10-15% (Sheatsley, 1948), sampling bias can become a serious problem (Hilgard and Payne, 1944). Heneman and Paterson (1949) reported reducing refusal rates to below 1%. In an unpublished survey by one of the writers, the refusal rate was cut from 15% to 4% by the use of a selected team of follow-up interviewers on initial refusals.

If it can be done, an appeal to the personal needs of a respondent, showing a way in which the survey can be of help to him, can be a highly effective inducer of rapport. However, if the motivation for cooperating is too personal, the respondent may distort his answers in the light of what he wants to get out of the interview.

Finally, the interviewer must attempt to make sure that the situation under which the interview is to occur will be optimal for rapport. If the interview is to be with a housewife, for example, and is to take a half to three-quarters of an hour, it would be wise to select a time of day when she can best spare that time, and a real attempt should be made to conduct the interview inside the house in a room where the interview is unlikely to be disturbed by others in the household. Not all questions in a study are affected by the presence of other people during an interview, but many are, and the pattern of role relationships (discussed above with reference to the respondent and interviewer) is made still more complex when others are present. In one study of the readjustment of returned veterans (Stouffer, *et al.*, 1949, pp. 596-644), it was found that when the veteran's wife was present during the interview, he was more likely to complain about unfair treatment by the public than if she were not present, and that he was less likely to express resentment toward civilians who asked him about his war experiences if others were present. Which set of statements more truly expresses the respondent's "real" attitudes is not so much the point here — the fact remains that the statements will be *different* if others are present, and comparability of cases is jeopardized.

Techniques of Interviewing Proper

Interviewing has often been called an art rather than a science, on the grounds that there are few scientifically established generalizations available concerning the usefulness of different interviewing techniques. Nevertheless, there are many points of technique which are widely agreed upon by practitioners, and it may be helpful to outline these here. Our summary is necessarily brief, and for more detailed discussions of interviewing technique the reader is referred to Bingham and Moore (1941), Rogers (1945), and Sheatsley (1951c).

The first question of technique concerns the interviewer's proper manner and bearing in asking the questions. Should he be responsive or detached? Respectful or man-to-man? Light or serious? Friendly or formal? In our earlier discussion of role, we pointed out the advantages of the interviewer's role as a stranger who is not involved in the respondent's normal system of personal interactions. This means that the interviewer should retain a certain amount of reserve; he should be dignified, and make it clear that he takes the interview seriously. At the same time, he should not be wooden: it is of the greatest importance for the interviewer to show genuine interest in what the respondent is saying. The interviewer must tread a fine line with respect to his responsiveness — it is a cardinal principle of interviewing that the interviewer must attempt not to reveal his own attitudes on the subject matter of the study: he must not show shock or disapproval over anything the respondent says, or nod enthusiastically when the respondent supports the interviewer's own point of view. If the interviewer is completely impassive and noncommittal, however, even when the respondent says something evidently intended to shock him, the natural conversational atmosphere of the interview will be destroyed, and rapport will suffer. The following has proved to be an effective compromise: the interviewer adopts a manner of friendly permissiveness. He laughs at the respondent's jokes, exclaims when the respondent says something evidently intended to be astonishing ("really?," "you don't say!"), makes supportive statements such as "I see your point," "that's understandable," "that's very interesting," and in other ways allows himself some of the emotional expressiveness which would be normal in the situation. He scrupulously avoids, however, direct approval or disapproval of the respondent's position — he never argues with the respondent, or says "I feel the same way myself."

Sometimes the respondent asks the interviewer where he stands on the issues under discussion. The interviewer may first simply try to deflect the question, by saying: "I'm interested in what *you* think about it," but if the question is pressed, he can either say that he hasn't made up his mind, or that he is not supposed to express his own opinion but would be glad to talk about it when the interview is over. (He will almost never be asked to make good on this offer!)

When conducting a structured interview, the interviewer must, of course, ask the questions word for word as they are provided for him. He can practice delivering them, however, until they sound and feel to him like his own words. Often after the interviewer has conducted a number of interviews using the same interview schedule, he begins to ask the questions too rapidly and in a "canned" fashion which makes it difficult for the respondent to understand him. He can usually overcome this tendency by pausing and looking at the respondent directly as he asks the question.

There is some difference of opinion in the field about how much time the respondent should be given to think over his answer. Kinsey believes that it is well to ask questions in a rapid-fire manner, so that the respondent will not have time to organize his defenses and think up an evasive answer. This technique is based upon the assumption that the truth is the first thing which enters the respondent's mind when he is asked a question. On the other hand, many researchers feel that permitting the respondent more time will reduce the number of "don't knows" and superficial or unconsidered answers. It would seem to the writers that for questions that are likely to meet with emotional resistance, the rapid-fire technique may be superior in some instances, but that, in most interviewing, allowing the respondent ample time to give a considered answer is preferable.

It has just been pointed out that in structured interviewing, the interviewer must ask the questions in the exact words of the schedule. But beyond this, in open-end interviewing, it is extremely important that the objectives of the study are met in each question. The interviewer must judge whether or not the question is understood by the respondent, and whether or not it is adequately answered. If it is not, probing must be done.

Users of open interviews agree that such probing should be nondirective. By this is meant that the question introduced by the interviewer should in no way influence the *direction* of the response. It should seek to elicit the desired information from the respondent without in any way making it likely that the respondent will lean more in one direction than another in responding to the probe or additional question. There has been considerable discussion by Rogers (1945) concerning the nature of nondirective probing in clinical interviewing and suggestions by him of how such probing should be used in research interviews. The Survey Research Center at the University of Michigan makes extensive use of such probing in its open-end interviews.

It should be noted that all nondirective probes do, in a sense, "direct" the respondent. Even if a respondent's statement is merely reflected back to him, or the question is repeated, or a neutral question is added, like "Tell me more about that" or "Why do you feel that way?" the mere fact that a particular point in the interview has been selected by the interviewer for probing will guide the respondent into a different channel than he would have chosen spontaneously. Indeed, the intent of introducing the probe is precisely to accomplish this. If the objective has not been met with the single use of the standard question, the interviewer probes in order to obtain information that will meet the objective. The question must be asked in such a way as to elicit a response which will provide information along a given dimension, and frequently the respondent must be "directed" to that dimension if the necessary data are to be obtained.

When an interviewer is employing a nondirective probe, he must learn to know when to stop probing. Often there is a reason why the respondent fails to understand the question or gives an answer which does not actually answer the question, namely, the question arouses anxiety in him. The interviewer must learn to detect signs of anxiety, and must judge when it is wisest to stop the line of questioning even if he has not obtained the information the objectives called for. To continue when the respondent is upset will jeopardize rapport for the entire interview, and it may be possible to return to the topic later when the respondent is more at ease. A danger is that the interviewer himself may feel anxious about a problem and may therefore be oversensitive to signs of disturbance in the respondent.

A common problem in interviewing occurs when the respondent gives information on a later question in the course of answering an early one. The safest thing to do is to ask the later question in its proper place, prefacing it by some remark such as: "You have already said something about this point, but . . ." and going on with the question. The advantage of this procedure is that the respondent may say quite a bit more that is highly relevant to the objective and still isn't likely to feel that the interviewer is being stupidly mechanical and not paying attention to his answers.

Related to this point is the problem of what to do if the respondent contradicts a point he has made earlier in the interview. This may arise in connection with factual information such as earnings or income, or it may arise in connection with an attitude. Sometimes the contradiction should be allowed to stand without comment in the interview: the existence of contradictory views in a single respondent may be an interesting finding of the study. But sometimes the study's objectives require that such contradictions be followed up in the interview. In such cases, the interviewer, in probing, can quote the respondent, but he must be careful not to antagonize him. One good device is for the interviewer to precede the reinstatement of the early remark with something like this: "I want to be sure to get this right, and I wonder if I haven't made a mistake. In my notes, I have . . ." This device saves the respondent's face and enables the interviewer to get the necessary information.

A problem sometimes encountered in the open interview is that of digression by the respondent. While in some interviewing situations the problem is one of getting beyond monosyllabic answers and encouraging the respondent to talk fully, in others the respondent is voluble and discursive, wanting to tell anecdotes or deliver lectures and failing to give material directly related to the subject of the interview. Serious problems are created when this occurs, for usually the interviewer has a list of topics or questions which he must cover in a limited time, and it will be impossible to complete the interview if digressions are not controlled. In her report of the problems of interviewing Russian Displaced Persons (among whom the problem of discursiveness is especially acute), Bauer (1953) has presented an excellent list of techniques for handling digressions by the respondent. She recommends that direct interruptions be avoided, since they can do irreparable damage to rapport. Instead, the following devices may be tried: (a) The "shunt" question — the respondent may be led back to relevant material by questions which relate what he has been saying to the topic in which the interviewer is interested. (b) Long personal anecdotes may be shortened by anticipating their ending: the interviewer, showing interest in the story, may guess how it came out, which will short-circuit some of the intermediate steps of the story. (c) Later digressions may be minimized if the interview is begun with a general question which gives the respondent an opportunity to bring out some of the things he especially wants to say. This is important if the respondent knows about the interview and its subject matter in advance, for he may have prepared a list of points he wants to cover with the interviewer. Often the things he wants to say are not the things which the interviewer particularly wants to hear, but it is necessary to give the respondent some opportunity to say what he has prepared, or he will show mounting irritation throughout the interview. Finally, one of the best protections against digression by the respondent is the precise formulation of questions.

RECORDING INTERVIEW DATA

A central problem of interview methodology is that of how the respondent's answer shall be recorded. Interviews have been recorded in the following way:

(1) *Write-up from memory.* The interviewer attempts to remember as much as possible of what the respondent has said, and makes notes on the content after the interview.

(2) *Field ratings.* The interviewer is provided with a set of field rating forms. He does not attempt to write down the respondent's remarks, but on the basis of the interview rates the respondent on the dimensions which the study is seeking to measure (war involvement, level of economic satisfaction, degree of democracy practiced in supervision, etc.). Each rating is usually based on a series of questions or on the entire protocol.

(3) *Field coding.* The interviewer codes specific answers to specific questions into predetermined categories, during the interview.

(4) *Verbatim notetaking.* While the respondent is talking, the interviewer attempts to take down his answer verbatim, or as nearly so as possible.

(5) *Mechanical recording.* The respondent talks into a microphone (which is either open or concealed). His remarks are recorded by means of a tape, wire, or disk recorder.

Recording from memory has the obvious disadvantage that it permits more distortion via the interviewer's biases than does immediate recording. Points in the interview which seemed dramatic or significant to the interviewer are likely to be sharpened, other points dropped out. Answers are likely to be recorded as being more consistent than they actually were. In short, all the forms of memory distortion which Bartlett (1932) has led us to expect should appear in this form of recording.

Apart from distortion, memory recording involves a considerable loss in the sheer *amount* of interview content that can be retained. Symonds and Dietrich (1941) made mechanical recordings of several fairly open interviews, and also had the same interviews written up from memory after varying periods of time by students who observed the interviews from behind a one-way screen. Using a system of counting points or "ideas" contained in the interviews, and weighting them according to their importance, the experimenters found that memory write-ups contained, on the average, 39% of the interview content if the write-up was done immediately after the interview, 30% after a lapse of 2 days, and 23% after a lapse of 7 days. Important points were recalled more fully than unimportant points, but even so, almost half the important points were omitted from memory write-ups.

Payne (1949) reports, in a comparison of mechanically recorded interviews with interviewer reports written from memory immediately after the interview, that 25% of the statements attributed by the interviewer to the respondent are clearly wrong. Covner (1944) finds, however, that while most of the items recorded by the interviewer after a guidance interview are accurate, the write-up is very greatly condensed (omitting important points) and introduces many changes into the temporal *sequence* of the interview. Points are brought into conjunction and made to seem related which were separated and unrelated in the actual interview.

The argument which has been advanced in favor of memory recording is that taking notes in the interview may upset the respondent, destroy his confidence in the anonymity of the interview, and produce an unnatural atmosphere. It is difficult to assess the damage, if any, that notetaking during an interview does to rapport. Experienced interviewers who do take notes report very few instances of the respondent's raising questions about the notes or asking him not to take notes. Occasionally, the respondent will want to reveal a bit of gossip to emphasize a point, and will say "Don't write this down, but . . ." or "off the record"; such remarks point to the fact that the interviewer's notebook sometimes makes the interview seem more formal, more official. This is not always detrimental to the interview, for it may mean that the respondent will give more consideration to his answers and not answer flippantly. Some respondents find it flattering that someone considers their views important enough to write them down, and in such cases, notetaking helps rapport. If the interviewer notices that his notebook seems to be making the respondent nervous, he can talk about notetaking a little before he begins the interview: "I hope you won't mind my taking some notes on what you say. I want to be sure I get it right" or "I talk to a good many people during the day, you know, and I don't want to get mixed up on who says what"; "I take notes on what people say, so I can refresh my memory later."

Actually, the majority of respondents seem to regard it as perfectly natural that the interviewer should take notes. If the interviewer is skillful, the notebook can be all but forgotten by the respondent as the interview progresses. Nearly all social scientists employing the interview feel that the advantages in accuracy of taking notes during an interview far outweigh any disadvantages, so that some form of recording during the interview itself is almost universal practice.

Field ratings are not very commonly used as a way of recording data. They may be practicable when the interviewers are very thoroughly familiar with the objectives of the study, and when the number of dimensions to be rated is relatively small. While little methodological work has been done to check the usefulness of field ratings, it is probable that they suffer from halo more than other forms of recording. Furthermore, it is difficult to maintain sufficiently similar frames of reference from one interviewer to another to ensure reliability. The Division of Program Surveys, U. S. Department of Agriculture, conducted a study during World War II of the effects of meat shortages on consumer food habits. Two sample cities were chosen, one of which had objectively experienced severe meat shortages, while the other had not. Interviewers were asked to rate each of their respondents on the severity of the meat shortages the respondent had experienced. The mean interviewer ratings were identical for the two cities. When coders in the home office rated the interviews on severity of shortage, however (without knowing which city the interviews were from), the mean ratings for the two cities were significantly different in the expected direction (Cannell, 1953). The interviewers, evidently, were using their own sample of cases as a frame of reference for "severity."

interviewer's write-up might be nearly identical with those based on a complete report. Much depends on what elements of the interview are to be used as cues for rating. If a rating is to be made of the respondent's anxiety level, for example, based upon such cues as blocking, over-attention to detail, evasiveness, and inconsistency, a complete interview record should be important, for interviewers would tend to drop out relevant elements in their notetaking. For less subtle ratings, the loss through notetaking may not be serious, although, of course, there is still considerable opportunity for interviewer bias to affect the study via the notetaking process. When interviewers are to attempt verbatim notes, it is important to train them to write up their interviews immediately. As they write, partial notes will recall fuller remarks the respondent made, but there is great danger that the interviewer will be reminded of the remarks of a *different* respondent if he has taken several other interviews between the taking and writing-up of a particular interview.

Experience is gradually accumulating with mechanical recording of interviews. Many technical difficulties remain to be ironed out — to the writers' knowledge, studies using mechanical recordings always suffer some loss of data due to inaudible recordings, but these losses need not be heavy. In a study by the Laboratory of Human Development at Harvard, which yielded 382 two-hour interviews conducted in respondents' homes, only 3 had to be discarded completely because of inaudible recording, and only 3 were partially lost. Technical improvements in mechanical recording devices are making their use more and more promising.

Mechanical recordings have been used both with and without the respondent's knowledge. There are difficult ethical questions involved here. But these questions aside, one may ask what are the effects of the respondent's knowing or not knowing a record is being made. The recording process has been concealed by some researchers on the grounds that it makes respondents uncomfortable, and that one can get more candid expressions of opinion when the respondent does not know he is speaking "for the record." On the other hand, if the recording is concealed, and if the respondent is told about it afterwards in order to obtain his permission to use the record, he may be irritated and feel that the interviewer has taken advantage of his good will. Of course, whatever damage to rapport this engenders occurs after the record has been obtained, but there may be serious public relations problems for further interviewing. There is reason, therefore, to tell the respondent and obtain his consent in advance whenever the subject matter of the study permits this, even when it is possible to conceal the recording device. In any case, in home interviewing with recorders, it is not usually possible to conceal the recorder. It may be fruitful, then, to consider how the mechanical recording process may be presented to the respondent so as to produce the least disturbance to the interview.

While little needs to be said to respondents about notetaking, since they usually take it for granted, mechanical recording equipment in a respondent's home is a different matter. It cannot be ignored, for it is usually bulky and often requires the respondent's help in getting it properly set up. A successful method of introducing recording equipment to respondents is to begin by asking where it could best be placed — the respondent's attention is then focused on mechanical problems, and the interviewer behaves as though recording is to be assumed, instead of asking permission to make a recording. The interviewer tests the machine by playing back a sentence or two of the respondent's words, and if there are children present, may allow them to hear their voices played back. Soon the respondent's curiosity is satisfied, and attention may be turned to the interview itself. With this procedure, no one in the previously mentioned study of 382 families objected to having a recording made.

A major problem with mechanical recording is the cost of having the interviews typed from the records. Skilled typists take 4 to 7 hours to type each hour of recorded interview, so that typing becomes at least as expensive as the interviewing itself. It has been suggested that ratings might be made directly from the recordings, so that typing would not be necessary. To the writers' knowledge, no comparison has yet been made of this procedure with that of rating from a typed transcript. One might expect that it would be very cumbersome to make over-all ratings directly from recordings, since the rater would have to refer back to materials scattered throughout the interview, and it is difficult to find the relevant sections of a recording. On the other hand, rating from recordings would permit the rater to take advantage of emphasis, intonation, pitch, and pauses as cues, which might be valuable for some purposes.

FIELD ADMINISTRATION

Any study involving large-scale interviewing necessitates the recruitment, training, and supervision of an interviewing staff. The number of interviewers may vary from just a few in a small local study to several hundred in large-scale national sample studies. Some research results have been reported on the validation of selection, training, and supervision methods. Stember and Hyman (1949c) report that although inexperienced interviewers show evidence of considerable error in the coding of interview material, experienced interviewers show little error. However, no alternative methods of selection and training were compared in this study. Other studies have investigated the effect of social class on interviewer performance (Katz, 1942).

A comprehensive study of interviewer characteristics as related to performance was reported by Sheatsley in a series of three articles (1950, 1951a, b). Three aspects of performance were rated on 1161 NORC interviewers: their handling of open-end questions, their paper work, and their carrying out of sampling instructions. An index based on these three sets of ratings was related to 19 interviewer characteristics such as experience, education, sex, political affiliation, etc. Although a few differences occurred which were significant at the .05 level, no differences of any important magnitude occurred. Sheatsley concludes that interviewer performance cannot be reliably predicted on the basis of the kind of information on interviewers' backgrounds and attitudes currently available. It may be of some interest to note that it is possible from Table 21 of the third Sheatsley article to discover the criteria that one of the outstanding survey organizations has used in the selection of its interviewing staff.

Most of the knowledge about this problem area derives from the experience of researchers rather than from research. Sheatsley (1951c) has presented an excellent summary of such accumulated wisdom. He lists desirable characteristics in the selection of interviewers in terms of adaptability, physical health, interest, intelligence, temperament, availability, appearance, and manner, and the absence of such special disqualifying characteristics as intense partisanship, the desire to regard the interview as a therapeutic device, and a long background of undesirable experience, such as selling. With respect to training and supervision, Sheatsley points out the need for inducing strong motivation in the staff concerning the nature and the importance of the study, and he stresses the need for careful training, including role playing and supervised practice interviews, and the continuation of careful encouragement and training while interviewing is in progress.

One problem of supervision is that of cheating, in which interviewers sometimes turn in reports without having conducted the interview or asked particular questions. Some systematic research has been reported on the incidence of such cheating behavior (Guest, 1947; Stewart and Flowerman, 1949), and spirited discussions of the problem were initiated in the literature by Crespi (1945-46).

Various methods are discussed for the detection of cheating: the use of postcards, re-interviews, phone calls to a subsample of respondents, and the use of internal interview checks. However, all discussants seem agreed that prevention is the most desirable single approach, and that the best steps are to ensure high interviewer morale through personal contact, interest in the study, high pay, high professional selection standards, well-designed questionnaires, and the like.

INTERVIEWING CHILDREN

A paramount problem in interviewing young children is that of learning the child's language and system of concepts, so as to permit phrasing questions which are meaningful to the child. It is extremely difficult for the adult interviewer to avoid assuming that the child can answer from an adult's frame of reference. For example, in interviewing adults, we frequently phrase questions in terms of what causes something else to occur. But, as Piaget (1929) has pointed out, the young child may not yet have developed a concept of causality, so that if we ask him such a question as "What makes the sun move?" the question (and consequently the answer) will not be meaningful. A study involving the interviewing of children normally requires a longer period of exploratory, unstructured interviewing and observation before it is possible to phrase good standardized questions. Piaget suggests that the research worker can find valuable clues to the nature of children's concepts in the questions which children ask spontaneously.

The question is sometimes asked: What is the youngest age at which children may be interviewed? The answer depends, of course, on the subject matter of the interview — for some purposes, the child can be interviewed almost as soon as he can talk. Radke (1946) used a list of 53 standardized open-end questions with children as young as three years old, covering such areas as affective relationship to parents, the child's views of the parents' techniques of discipline, and the child's ideas of what constitutes "good" and "bad" behavior. She found that although the thought processes of the younger children were slower, they could give perfectly meaningful answers.

As in any interviewing, the interviewer who works with children must devote considerable attention to getting good rapport. The child may be more reluctant than adults to express his own thoughts and feelings because (1) he assumes that everything he knows is already known by adults, and (2) he is afraid of saying something wrong or foolish (Piaget, 1929). Probably one of the reasons it is so helpful to play with a child before starting to ask him questions, and to put the questioning in the context of a game, is that these procedures help to free the child from the fear that his answers are to be judged right or wrong.

A further important element of good rapport with children is the proper timing of the introduction of the more formal parts of the research process. When he is first brought into the research situation, the child will want to explore and satisfy his curiosity about the room and the objects in it. He must be given time to begin to feel at home, but must not have time enough to become bored. A good way to determine when to begin a structured interview is to allow the child to give the signal himself: "We want to play a game together in a minute. Will you tell me when you're ready to begin?"

The fact that the child is unsophisticated about the meaning of his answers to the adult interviewer means that he will not try to cover up or disguise certain kinds of answers as much as an adult will. For example, a child will usually tell freely and openly about his dreams, whereas many adults have learned that adverse personality judgments may be made on the basis of their dreams and will be careful about telling them. Furthermore, a child will be less troubled by the desire to be consistent, and will usually express inconsistent beliefs without embarrassment. This is not to say that young children are free of ego defenses in an interview. If they are being asked about areas which are anxiety-producing (for example, their own "naughty" behavior), defensive reactions may emerge, such as leaving the field ("I don't want to play any more"), failing to understand the question, changing the subject, nonsense answers, etc. To prevent these reactions, the interviewer must be as permissive as possible, making it clear that he will not disapprove of the child or punish him for anything he may say, and must phrase the questions carefully. One interviewer found, for example, that he could not ask children "When you have done something you shouldn't, do you tell your mommy or do you hide it from her?" Instead he had to use the more euphemistic expression: "Do you tell your mommy or do you keep it a secret?" Keeping something a secret was evidently more acceptable to five-year-olds than hiding something from their mothers. If a child shows any signs of disturbance over a question, the interviewer often finds it useful to interrupt the interview for a period of play or refreshments, until the atmosphere is lightened.

Piaget feels that a major problem in interviewing children is their suggestibility. They are accustomed to taking cues from adults when they are unsure about what they should say or do, and they are therefore quick to agree with the interviewer if he says "I imagine you feel (think) thus and so, don't you?" The usual rules about nondirective phrasing of questions are, then, even more important for child interviews than adult interviews. As a way of checking on whether a child's statement is merely a response to a suggestion from the interviewer, Piaget suggests that the research worker should use counter-suggestion, and ask about the same subject in different ways at different times. The suggested belief should disappear rapidly.

A further problem Piaget mentions in connection with interviewing young children is what he calls "perseveration." If the interviewer asks in order whether the sun, wind, clouds, rain, etc., are alive, the child is likely to say "yes" to the entire series if he says "yes" to the first one. This same tendency to complete series is seen in doll play, where the child seems to feel that every member of the doll family must be put through the same act sequence before a new act can be introduced. The interviewer can guard against this excessive stereotypy by breaking the sequence of questions and coming back to a topic after a new set has intervened.

The young child is, of course, not very skilled in the use of words, and he cannot continue long with a situation which takes an exclusively verbal form any more than he can carry on a sustained conversation. It has been found very helpful to combine interviewing with play materials, so that the child may act out some of his answers as well as verbalize them. The doll-play interview (Levy, 1933; Conn, 1938; Ammons and Ammons, 1952) has proved very useful. In this technique, the dolls are asked questions, and the child answers for the dolls. The child may have the doll talk, or act, or both. Extremely high observation and scoring reliabilities are reported for this technique (Ammons), as well as good session-to-session reliabilities. Ammons and Ammons found that children were willing and able to answer the experimenter's questions more often when the questions were phrased in terms of what the doll would *do* in a given situation, rather than what the doll would *say* — a further indication of the limitations upon the child's ability to deal with exclusively verbal material.

INTERVIEWING IN OTHER SOCIETIES

The experiences of field workers in other societies are of especial interest because they highlight some of the methodological problems the social scientist encounters (but sometimes overlooks) when he attempts to work with people from different ethnic backgrounds or different socioeconomic classes within his own culture. Furthermore, as social psychologists become increasingly interested in the fields of "culture and personality" and "national character," and in cross-cultural tests of psychological hypotheses, they must concern themselves with the problems of gathering data in societies other than their own.

Social anthropology is rich in experience with the interview as a tool for work in other societies. Benjamin Paul (1953) has provided an excellent summary of the literature on anthropological interviewing, and offers valuable insight into the special problem of role taking. In comparing social-psychological with anthropological technique, it is important to note that anthropologists (especially those working in nonliterate societies) have worked primarily with a small number of *informants* in the societies they have studied; they have seldom attempted to do large-scale interviewing with a cross section of respondents, and have tended not to be concerned with standardizing measuring instruments. The interviews with informants are generally very long, taking many sessions, and interview technique varies from very free, nondirective interviewing to interviewing in which the ethnographer takes an active part in guiding the informant along chosen topical paths.

One of the first major problems encountered by the field worker in another society is how to present himself to that society. The role of scientist — someone who is trying to find out things simply to accumulate knowledge — is familiar to Western societies and is generally regarded with sympathy, but this role is completely unintelligible to some peoples. Especially where most people are not literate, the spectacle of someone with a notebook and a typewriter asking questions and making notes can be frightening. It is of the greatest importance to be introduced by someone who is trusted by potential informants, and to attempt to make one's purpose (or part of it) clear and acceptable. The field worker is sometimes helped by the fact that he is a foreigner, and strange behavior may be tolerated in him when it would not be in an in-group member. Essentially, the field worker must be *accepted* in some role before he can begin interviewing, and gaining acceptance may be a slow, arduous process. In a village in Afghanistan, for example, a field worker had to wrestle publicly with the village wrestling champion and take high hurdles on horseback before he was accepted.

A second major problem is that of language. Some field work has been done exclusively through interpreters. In such work, of course, the information received is filtered through the personality and status of the interpreter. How important this can be is apparent, Kluckhohn (1945) points out, when the field worker is forced to change interpreters in the midst of his work with a particular informant. After the change, a different sort of material is likely to be forthcoming, either because the new interpreter censors different things, or because of the informant's tendency to put things in a different light before this new person. When the ethnographer wishes to study topics which are surrounded by taboos and overlaid with anxiety, it is especially dangerous to obtain all his information through an interpreter; any advantage that the ethnographer might have in receiving confidences because of his being a stranger is lost.

Field workers have often worked with a "contact" language — a language which is secondary both to the field worker and to the informant, but which permits direct conversation. Pidgin English has often been used for this purpose. As Margaret Mead (1939, pp. 193-194) cogently points out, contact languages have generally grown up in the context of special kinds of relationships between groups: the master-servant relationship, the conqueror-conquered relationship, etc. She says: "These contact languages often have a complete ethos associated with them, and are regarded by one or both of the cultures in contact as instruments of hypocrisy, of wheedling, ingratiating, or bullying, and a great many values besides those specifically linguistic may be jeopardized (by their use.) " It is evident that field work benefits enormously if the field worker can learn the native language of the group he is studying, although to do so involves an enormous burden of time and effort.

There are several advantages to beginning field work with a period devoted to learning the language. It is easy for the people in the society to understand what the field worker is doing when he is learning a language, and they are usually glad to help. Furthermore, the field worker learns not only the denotative value of words, but a great deal about the way in which the people think — the way they categorize phenomena. Such knowledge is indispensable when it comes to phrasing questions in such a way that they will be meaningful to the respondents. He may find, for example, that there are few words referring to time periods, and from this observation he may be led to the discovery that the society has a different time orientation than our own, so that questions phrased in terms such as "How long ago was it when . . ." would be difficult for the respondent to answer.

In fact, before a field worker can do intelligent interviewing on particular topics, he must steep himself in the culture to the fullest possible extent. In interviewing in our own society, many shared customs are implicit in the questions we ask (we take for granted, for example, the fact that every seventh day is a holiday, or that the group of people sharing a dwelling is likely to be made up of a husband, wife, and children). The field worker in another society must start very far back — not only will he be ignorant of the shared customs and beliefs that everyone takes for granted, but he won't know how to ask about them. He must therefore spend a much larger proportion of his time in the exploratory and "scouting" phases of his research than would be necessary to study a specific topic in his own society. There is a certain dilemma here: even the exploratory phase will involve interviewing, and how shall questions be phrased? The best solution seems to be that they should be as open as possible. They can be phrased in the form: "Tell me about . . ." and should be followed up with nondirective probes. One pitfall of the beginning interviewer in another society is that he will want to appear intelligent and understanding, so as not to invite the contempt of the people with whom he is working. He is therefore likely to nod his head and say "Oh yes, I understand" when he doesn't understand or understands only in a completely alien frame of reference. He must learn to say easily: "I don't quite understand that. Would you tell me more about it?"

Some field workers have used pictures to great advantage in the initial, most open stages of their interviewing. They take a picture of a village scene, and then ask a number of different people what the people in the picture are doing, what they are about to do, how they feel, etc. This technique helps the interviewer to get the "feel" of the society with very little directive questioning on his part.

Perhaps the most common fault of the novice interviewer, in this as in so many other settings, is to put words into the mouth of the respondent. For example, in interviewing about courtship patterns, one novice interviewer was told that a young man had told his father he was interested in a particular girl. The interviewer said: "And then your father went to see her father, is that right?" This kind of interviewing is particularly dangerous in other societies, where the interviewer's guesses about what the respondent has in mind are likely to be inaccurate, and where the respondent may be accustomed to showing deference toward people like the interviewer, so that he feels constrained to agree with what the interviewer says.

As in any interviewing, the field worker in another society must first obtain the interest and cooperation of his respondents. To do so he needs to know a good deal about what motivates people in the society. Some sort of economic reward is often helpful in enlisting cooperation (Kluckhohn, 1945). The feeling of being misunderstood in one's own group and the desire to talk to someone who will be sympathetic is often sufficient motivation for cooperation, but this motive must be tapped with care, for it is likely to draw deviants, and the field worker is usually more interested in the individual who is in some sense "typical" of the people in the society. Sometimes an informant will be motivated by pride in his own group. For example, Indian mothers became interested in being interviewed about their child-rearing methods when the field worker asked them how they succeeded in training their children to be so quiet and well-behaved — pointing out the fact that white children cry much more (Chasdi, 1952).

Good rapport requires, of course, that the field worker conduct himself according to the rules of politeness prescribed by the society he is studying. These rules may be very different from those under which he has been brought up. For example, among the Navaho, it is impolite to look someone in the eye — the polite interviewer lowers his eyes when introduced, and then sits *beside* his respondent, and both gaze off into the distance while talking. Or, for another example, in some societies it is impolite, while talking to one member of a society, to refer to another member by name.

During his exploratory work, the interviewer usually needs to discover something about what topics are taboo in the society. To obtain such information, he will find it advantageous to begin his interviewing with a particular respondent with topics which are emotionally neutral, and turn gradually to the more touchy subjects as rapport develops. In general, language, genealogy, and material culture are good topics to start with. Such topics as religion, status relationships, and attitudes toward sex and aggression should come later.

As in any interviewing, the ethnographer must watch for signs of anxiety in the respondent, and adapt himself accordingly. In other societies the signs of anxiety may not be the same as they are in his own; for example, instead of becoming restless and fidgety under anxiety, the respondent may become stonily impassive, and the interviewer must learn to recognize the meaning of this behavior.

The mounting interest in cross-cultural studies raises the question of the feasibility of greater standardization in field work methods from one society to another. Also, whenever the scientist is interested in comparing individuals within a society (in studying individual differences), he must consider the advantages of standardizing his interviewing techniques from one respondent to another. Lists of standardized questions with large numbers of respondents have been used within a nonliterate society (F. Kluckhohn, *et al.*, 1953), and the authors' own experience in conducting interview studies in two societies other than our own (both literate and both Westernized, however) suggest that it is entirely feasible to standardize questions and obtain comparable information from several societies. Of course, the problem of *defining* underlying variables which will be valid cross-culturally is very complex. For example, suppose one wished to compare the amount of aggression shown in a number of societies, or compare the severity with which children are punished for aggression. In one society, aggression is shown by making a doll in the image of the enemy and destroying it; in another, the aggressor may give the enemy a white feather; in another, he may spit at his feet; in another, he may "damn with faint praise." How to decide what will be included as aggression for purposes of cross-cultural comparison? For other kinds of problems, getting cross-cultural comparability of variables is not nearly so difficult, and simple translation of questions from one language to another may be in order. When the research worker wishes to translate questions into a language which is unfamiliar to him, incidentally, one technique which has been used successfully is to have the questions translated by one bilingual person, and then have them translated back into the original language by a second bilingual person. The differences between the research worker's original form and the retranslated version will reveal to him any serious alterations in meaning which grew out of the translation process.

SOURCES OF INTERVIEWER ERROR

Of all the methodological problems involved in interviewing, the one which has been most fully studied is the problem of interviewer error (bias). It has been customary to use the term *interviewer bias* for many of the kinds of errors which will be examined in this section. Properly speaking, however, the word *bias* refers only to systematic or nonrandom errors. Since many of the sources of interviewer effects on interview data do not tend to shift the data in a particular direction but rather are more or less randomly distributed, we prefer to use the more inclusive term *interviewer error.* No attempt can be made here to list all the studies on interviewer error. Before we attempt to summarize some of the findings in this area, a brief discussion of the meaning of the term may be in order.

Interest in the problem grew out of the observation that the set of answers reported by one interviewer would differ from the set reported by another interviewer, even when the samples were comparable. Furthermore, it appeared that some interviewers obtained answers which were similar to their own views, and the suspicion naturally arose that the interviewer had somehow influenced the respondent. In a study where absolute levels of response are important (how many people expect to vote for a certain candidate, how many own cars, etc.), it would obviously be damaging to the study if all the interviewers were influencing their respondents in the same direction. Even if the influence of various interviewers counterbalanced each other, error would be introduced into the measurement process, and intercorrelation among variables would be seriously affected.

How to measure interviewer error? The simplest approach would be to obtain some independent measure against which to check the interviewer's data. Thus, interviewer reports of whether a set of respondents voted could be checked against poll records. But discrepancies between outside records and interview reports might be due to poor question wording, or poor memory on the part of the respondents, rather than to any error by the interviewer. Furthermore, when a study is concerned with attitudes, there is no factual base line against which to compare the respondent's answers. Suppose, for example, that a respondent expresses more liberal attitudes on race to a Negro interviewer than to a white interviewer. Which interviewer is biasing the respondent? Which statement represents his "true" attitude? Clearly, it is possible that the respondent actually has different attitudes in the two different social contexts — that his attitudes are situationally determined. The interview must set *some* social situation for the expression of attitudes, and the attitudes expressed will necessarily be a function of that situation. The problem of measurement arises when the interviewers vary in their appearance, their manner, their words, so that the measurement situation is not comparable from one case to another.

The amount of interviewer error, as currently conceived, means the amount of variance in data which may be attributed to interviewers rather than respondents. It is studied by comparing sets of answers which different interviewers obtain from similar samples, to see whether the distributions show more difference than would be expected by chance.

Six possible sources of error in interviewer performance and representative research concerning these will be discussed: (1) the appearance and manner of the interviewer, (2) the way in which interviewers word or deliver questions, (3) the interviewer's own attitudes, (4) the interviewer's expectations of the respondent's attitudes, (5) interviewer variations in the fullness of probing, and (6) interviewer variations in the recording of respondents' answers.

(1) *Appearance and manner of interviewer.* Katz's (1942) study showed important differences in the responses of persons interviewed by working class and white collar interviewers. In a study conducted by OPOR in 1942 (Cantril, 1944), Negro respondents were questioned on their attitudes towards German or Japanese treatment of American Negroes in the event of their winning the war. Negro interviewers obtained a far higher proportion of pro-Axis sentiments than did white ones. Sampling differences, namely, the particular respondents selected by white and Negro interviewers, could account for these differences. However, a Research Branch Study (Stouffer, *et al.*, 1950, pp. 718-721) with a similar design in which it is less likely that respondent selection could have caused the differences reports comparable results. It is possible that other factors could account for the obtained differences. Such factors as cheating or failure to probe, particularly if such factors are correlated with class membership or race, might be partly responsible. However, the results are consistent with the hypothesis that the appearance and manner of the interviewer are factors biasing interview data. Most survey researchers feel that the manner and appearance of interviewers is important, and they often attempt to achieve inconspicuous and relatively uniform dress among their interviewers, in spite of the fact that not much definitive research has been done on this problem.

(2) *Question wording by interviewers.* In previous sections of this chapter, the effects of different ways of wording questions both in standardized and in unstandardized interviewing have been discussed. In unstandardized interviews, the interviewer has great latitude in question wording, and systematic differences in the data due to interviewer selection of question wording are clearly possible. Even in standardized interviewing this factor can be important. This is particularly possible when probing is encouraged or permitted, either with open-end or closed-end questioning. In studies reported by Guest (1947) and by Stewart and Flowerman (1949), in which interviews were recorded without the knowledge of the interviewers, considerable differences among interviewers occurred. In Guest's study, both closed- and open-end questions were used. The closed-end questions theoretically should provide no opportunity for interviewers to manipulate question wording. Nevertheless, he found word changes of both "major" and "minor"

sorts among his interviews. Furthermore, in the use of probes designed to obtain specific or complete answers to his open-end questions, he found many major differences among interviewers.

Stewart and Flowerman found many instances where questions were not asked at all, although answers were recorded even by interviewers who were highly conscientious in seeking out their assigned respondents. Both studies reported that the usual quality control procedures in examining the interviewer's written records failed to reveal the nature and frequency of such practices. The latter investigators developed an office procedure they call the "displacement measure" which correlated highly with the errors revealed on the interview recordings. It has been shown that a scaling procedure such as the Guttman scale can also reveal the presence of such interviewer cheating (Elinson and Cisin, 1948).

(3) *Effects of interviewers' own attitudes.* Public opinion researchers have evinced considerable interest in the problem of the possible biasing effects of the opinions the interviewers themselves have in the attitude areas being studied. Does a Republican interviewer tend to obtain more or fewer Republican opinions from his respondents than a Democratic one? Does an internationalist-minded interviewer get more or fewer internationalist attitudes from respondents than does an interviewer who is more isolationist or nationally minded? Typical of the studies in this area is one reported in Cantril (1944). The study was a pre-election poll being made prior to the 1940 presidential elections. Each interviewer was asked to fill out the same ballot as the respondents were given in a national sample poll. The interviewers then returned this ballot with their completed interviews. Then the distributions of the interviews taken by interviewers who had a given attitude were compared with those obtained by interviewers with opposite attitudes. Interviewers who favored helping England obtained one and one-half times as many respondents favoring helping England compared with those favoring keeping out of war at all costs. On the other hand, interviewers who felt that we should keep out of war almost reversed this figure for their sample. Similar differences occurred on presidential choices. Since quota sampling methods were employed in both these studies, it is not possible to determine to what extent respondent selection is responsible for these differences and to what extent, if any, interviewers were affecting the expressed attitudes of their respondents.

It is also interesting to note that for interviewers working in large cities, no such biasing effects occurred. The author suggests that it may be more difficult for interviewers to select in advance, either consciously or unconsciously, respondents whose attitudes will correspond to his own in large communities, where fewer people are known to him.

Katz's study, reported above, also throws considerable light on this problem. He found that working-class interviewers tended more often to select middle-class respondents than did middle-class interviewers. However, the working-class interviewers tended to elicit more pro-working-class attitudes from middle-class respondents. Even in this study, however, it is possible that working-class interviewers tended to select middle-class respondents whose views were more like their own than was the case for middle-class interviewers. Therefore, there is no clear demonstration of effects of interviewers on the expressed attitudes of the respondents.

It remained for the comprehensive study of Feldman, Hyman, and Hart (1951-52) to separate out the relative contributions of respondent selection and interviewer effects on respondents' reports. In their experiment, nine interviewers were assigned to each of five segments of Denver. Respondents were randomly preselected for each interviewer so that samples were comparable from interviewer to interviewer within each sector, thus controlling the possible biasing effects of respondent selection. A predesigned schedule and uniform instructions and supervision were employed. Whereas previous studies had shown biasing effects when very few interviewers were being compared (Rice, 1929; Shapiro and Eberhart, 1947), this study, providing five replications with nine interviewers each, found relatively little bias. Significant differences were found among individual interviewers on but seven of the 21 opinion closed-end questions employed and on but one out of seven factual questions. Some rather striking differences did occur, however, on two kinds of items: field ratings or over-all codes of respondents by interviewers showed wide differences, as did the type and quantity of information obtained in open-end questions. These latter differences were not found to be systematically related to the interviewer's own attitudes, but rather to his expectations of the respondents' attitudes. Further evidence on this point is presented in (4) below.

It should be noted that, in this experiment, regular experienced interviewers were employed in much the same way they would be in a typical professionally conducted opinion-attitude survey. These studies indicate that when interviewers are free to select their respondents, as in quota sampling, very considerable differences can occur in the distributions of responses of allegedly comparable samples of respondents. However, few such differences have been found to occur due to the influence of experienced, well-trained interviewers on the respondents' expressed attitudes.

(4) *Effects of interviewers' expectations.* Two published studies (Smith and Hyman, 1950; Wyatt and Campbell, 1950) suggest that interviewers' expectations of the respondents' attitudes have a greater influence on how the interviewer will perceive and record the opinion of the respondent than the interviewer's own attitude. Smith and Hyman used the device of a recorded interview in which a "planted" respondent evinced an attitude not generally correlated with the other attitudes he revealed. For instance, the respondent would generally take a strong internationalist position on foreign affairs, and then suddenly, in answer to a particular question, take an isolationist position. Interviewers were also polled on their own attitudes. When the respondent took an unexpectedly isolationist position, no differences occurred between "isolationist" and "interventionist" interviewers. Only one-fifth of each group, however, coded the inconsistent isolationist response correctly. When the respondent took an expected interventionist position, many more of both groups of interviewers were correct: 78% of the interventionist interviewers and 61% of the isolationist interviewers. Wyatt and Campbell found no significant biasing effects due to either source, but the tendency towards bias was more pronounced due to interviewers' predictions of population distributions than due to the interviewers' own opinions.

In a recent and as yet unpublished study, in which highly trained interviewers and a carefully designed questionnaire were used, Cannell (1953) found no biasing effects of interviewers' expectations. Apparently both sources of bias, interviewer attitudes and expectations, can be reduced to a minimum under appropriate conditions.

(5) *Interviewer variations in probing.* In open-end interviewing, whether the interviewer attempts to record the complete answers of his respondents or codes them directly into precoded categories while the interview is in progress, variations can occur in the extent to which he will continue to question a respondent if the original question is not adequately answered. The extent to which a respondent is pressed to state an opinion may differ from interviewer to interviewer and even from respondent to respondent with the same interviewer. Thus, in the excellent large-scale study of interviewer effects on survey data reported above (Feldman, Hyman, and Hart, 1951-52), wide differences occurred in the numbers of answers obtained by different interviewers assigned comparable samples. Some interviewers evidently probe considerably more than others and thus obtain considerably more information. Shapiro and Eberhart (1947) also obtained similar results.

(6) *Interviewer variations in recording.* The problems involved in interviewer recording of respondents' answers were discussed in some detail in the preceding section on recording interviews. There are several studies relevant to this problem. In an investigation by Guest (1947), a wire recorder was concealed in a respondent's home. Data are reported on 15 college student interviewers who interviewed this one respondent. Questions of both the open-end and multiple-choice type were used. Some of these were informational and others attitudinal. The subject matter dealt with attitudes towards psychologists. There were thirty-two errors of commission involving recording something either not said or said differently, and sixty errors of omission in not recording something said. In addition, 41 cases of not differentiating probes from regular questions in the write-up were discovered. Unfortunately, the proportion of these errors to total possible errors is not reported, and it would also be of interest to know just what changes in tables of results these errors would cause. In any case, the frequency and character of the recording errors reported suggest that the problem of interviewer recording bias may well be a serious one.

A second experiment on this problem is reported by Stewart and Flowerman (1949). The authors obtained full wire recordings of 33 interviews involving 15 interviewers. These interviewers were trained on a fifty-item questionnaire dealing with attitudes towards minority groups. The 33 interviews recorded were with different respondents who were trained to give the interviewers as difficult a time as possible. As in the previous study, considerable error occurred in the way interviewers recorded answers and, also as in the previous study, the usual methods of interviewer evaluation failed to predict which interviewers made the most errors.

Both these studies, together with the considerations listed in the previous section, suggest that interviewer recording error is indeed a serious problem in the use of interview data.

The work on interviewer error to date points up the seriousness of this problem in the use of the interview as a research tool. Some of the research cited above is particularly helpful in that it points to ways of overcoming interviewer error.

VALIDITY OF INTERVIEW DATA

As in the case of any measure, it is essential to determine the extent to which the interview constitutes a valid measuring instrument. The question of the validity of "the interview" in general is too broad. Validity depends upon the nature of the particular interview technique, the subject matter of the study, and the circumstances in which the interview occurs. But research workers using the interview must face the question of whether answers in an interview have any relation to actual behavior, the underlying attitudes which govern behavior, or the factual events concerning which they are asked to report. We know from studies on the reliability of witnesses that many people are very poor reporters and will misreport even if they are asked about recent events. Furthermore, with the passage of time, omissions and distortions become much more pronounced. And we know that respondents will want to put themselves in a favorable light before the interviewer, and may therefore distort or conceal (either consciously or unconsciously) their attitudes and behavior. If these things are true, how can we rely upon an interview to give us valid data? It is not enough that those who use interviewing as a method usually have a strong conviction that their respondents are giving them meaningful answers — the answers "ring true," they are colorful, and it seems impossible that respondents could invent out of whole cloth the experiences and feelings they report. There should be controlled empirical evidence. McNemar (1946) pointed out that few validity studies had been reported. What is the current evidence concerning interview validity?

There are several ways of stating the validity problem. In attitude measurement, the problem is often described in terms of what predictions can be made about how a respondent will behave, say in the voting booth, or in the presence of his immediate supervisor, or towards his child, on the basis of information obtained from interviewing him. Another way of stating the validity problem is perhaps of no lesser importance. That is, to what extent are interview data measures of what they purport to measure? An attitude, for example, towards continuing economic aid to Western Europe need have no particular behavioral expression other than a verbal one. The question here is not so much "What acts do interview data on such an attitude predict?" as "To what extent do such data accurately reflect a respondent's underlying attitude on this issue?"

Before proceeding to the discussion of varieties of validity studies, it might be well to point out the distinction between group validation and individual validation as applied to interview studies. In group validation, predictions from interview data are made as to how a population will behave. For example, in a pre-election poll, predictions may be made concerning what proportions of the electorate will vote for each candidate. An interesting example of a method of group validation is provided by the Survey Research Center in its studies of consumer finances. The total income, government bond holdings, or bank deposits reported by a cross-section national sample of spending units, is multiplied by the reciprocal of the sampling ratio. (That is, if every 10,000th spending unit is interviewed, the total income of the respondents in the sample would be multiplied by 10,000.) The resulting figures should represent the total national income of spending units, or the total individual holdings in Government bonds or bank deposits. The aggregate figures thus derived are compared with Federal Reserve Board estimates of these quantities based on their reports from banks and other financial sources. Thus, even though no outside data are available on the *distribution* of incomes or savings, outside aggregates may be used to check the validity of a survey's over-all findings.

In individual validations, a given individual's behavior may be predicted on the basis of interview data, and his subsequent behavior can be observed and compared with the prediction. For example, in the Erie and Elmira county panel studies of the determinants of voting behavior, the same individual is followed through a campaign and his actual voting behavior is determined by interview after the election. Similarly, in a Survey Research Center economic behavior study, check-ups were made the following year to determine how many of the individuals who reported plans to purchase durable items during the previous interview actually did so.

Either individual or group comparisons may be made in any of the four types of validation procedures discussed below. Four different varieties of validity data have been collected, bearing upon the above statements of the validity problem.

(1) *Comparison of respondents' factual reports with outside reports or objective evidence of the fact under consideration.* Respondents are asked, "Did you vote in 1948?" "What magazines do you take?" "Do you own a car?" "What was your income last year?" "How often do you spank your child?" etc., and the answers are checked against other evidence.

Hyman (1944-45) reports three such studies based on wartime interviewing. A list of 243 respondents who had redeemed U. S. War Savings Bonds were interviewed, and 42 (17%) of these denied having redeemed bonds. Of the grocers who, when interviewed, reported having a certain government poster on display, 42% did not actually have it on display, and 4% of workers who had been absent from work said they had not been absent. (The 4% figure was obtained on open-end interviews. When a closed-end question was used, the distortion was 27%.) These studies suggest some of the most valuable ways of obtaining valid criterion information. In all of the situations above, the distortion was in the direction of social approval. In the absentee situation, the open-end question was least objectionably worded from this standpoint.

A recent study reported by Parry and Crossley (1950) presents considerable data on the validity of interview data on eight classes of factual information. It is interesting to note in this study that certain information is subject to far less response error than other information. When the respondent is asked for factual information that he is likely to know about in precise detail, and when the correct reporting of this information is not likely to subject him to feelings of guilt, he is apparently able to report accurately. For example, all but 3% of those from whom the information was obtained correctly reported whether or not they owned a car, 96% reported accurately that they either did or did not own a house, 98% that they did or did not have a telephone. When, however, the respondent is asked for information in which memory errors are likely to occur or in which guilt feelings are likely to arise when answering in a certain way, response error does tend to occur. For example, the more recent the presidential election, the more likely the respondent is to report accurately about his voting. In 1949, 86% reported correctly whether or not they had voted in 1948; this figure dropped to 73% for the 1944 elections; and in both cases the error is predominantly in the direction of voting reported by those who did not actually vote. On community chest contributions, where social pressure to contribute is strong and where the subjects are not likely to be aware of any possibility of verifying their answers, 34% reported giving who did not give. The corresponding figure for those who actually contributed but reported not doing so is unfortunately not given. Even if allowance is made for some error in the criterion listings, the response error is no doubt substantial. It is at least clear from this study that the correct reporting by interview of factual information cannot always be safely assumed, and that certain kinds of information appear subject to considerable response error.

Kinsey, Pomeroy, and Martin (1948) have devoted considerable attention to the validity problem in their well-known studies of human sexual behavior. One of their methods is to compare the data obtained from husbands with that from their wives. They found that *incidence* of certain behavior patterns was more often independently agreed upon by both members of a marriage than was the *frequency* of such behavior. Kinsey and colleagues also found that except for the aged, recent events were recalled more easily than remote ones. They also compared interview data (obtained from prison inmates) on homosexual *incidents* with an independent estimate by an observer of the number and identity of men believed to be active homosexuals, and found about 85% agreement.

The individual validation of questionnaire data against independent self-observation of behavior has been done by Goodenough (1931). She asked parents to check from a list of methods for controlling anger in children those which they employed. Following this, she had them keep a diary on what they did each time the child got angry. Little correlation was found between the sets of data. It may well be the case, as is so often true in our society, that notions of ideal behavior were being checked and that actual practice fell considerably short of these ideals.

Radke (1946) also attempted to validate parental reports of child-rearing behavior by obtaining independent measures from each of the two parents and also from parents and children of the same families. The correlation between parents was .64 and for parent and child, .42. The parents represented themselves to be warmer, less punitive, and more consistent than did their children.

(2) *Prediction of future behavior on the*

basis of respondents' statements of their plans and expectations. By far the most extensive source of such data is in the area of pre-election forecasting. There are, however, a number of reasons why these are not completely satisfactory as validity checks.

Sampling problems, for example, are especially complex in election forecasting. In order to interview a representative sample of voters, the pollster must be able to specify, first of all, the population who will actually vote on election day. Secondly, the campaign does not cease from the time interviewing is begun until election day, and presumably many people do change their minds in the course of at least some campaigns. In other words, even if pre-election polls were completely valid measures of the voting intentions of the population sampled, it would not necessarily follow that the group predictions based on these data would be accurate for election day (SSRC Comm., 1948). When these and other factors, such as the undecided vote, are taken into account, the remarkable fact usually is how closely pre-election poll *data* do predict aggregate percentage votes. In addition to group validation procedures available through election returns, various additional checks are available through comparing the predicted behavior of subgroups such as the working class or the upper class with the election statistics of precincts or wards which are known to be heavily populated by the members of a class.

Students of economic behavior are often concerned with predicting the trends of consumer purchasing and saving, for these are closely related to the economic cycle. Interview studies have been done in which consumers are asked their buying and saving plans and, from their replies, an effort is made to predict economic trends in the near future (Katona, 1951). In several instances, trends predicted from such surveys have proved to be accurate, but much more experience is needed before the method can be considered fully tested.

(3) *Consistency of statements relating to the same underlying attitude.* Internal consistency measures are usually considered solely as reliability measures. For example, if the same question is asked in somewhat different form at two different points in the interview and the answers to the two questions are highly correlated, the measure may be considered a reliable one. However, suppose that a number of questions are asked which differ not only in form but contain substantively different subject matter, these different elements of subject matter being related to each other in a predicted psychologically meaningful or logical fashion. High correlations among these elements suggest that these questions are all tapping a common attitude or set of attitudes and are measuring what they purport to measure. If, for example, a respondent indicates through one question that he believes in sharing all our military secrets with Britain and in answer to another question says we should standardize our military equipment with that of the British and on still a third question that U. S. and British armies should be integrated at the Divisional level, the evidence begins to mount up that that respondent believes that there ought to be a close military alliance between Britain and the U. S. Similarly, when a group of items forms a Guttman scale with a reproducibility score in the nineties, not only is high reliability being indicated, but the fact that the attitudes expressed by the replies to the series of questions have psychological consistency suggests that the scale is measuring an attitude that can be labeled with some confidence in terms of its verbally expressed content.

(4) *Indirect validation: correlation of interview data with data which have a predictable relationship.* This form of validation is perhaps more commonly employed than any other (although it is often not thought of explicitly as a form of validation). On the basis of theory and previous research, the social scientist predicts that a variable which he measures through the interview ought to be related to another variable (to be measured through interviewing or in some other way). If the relationship actually emerges from the study, there is presumptive evidence that the interview measure was valid. Of course, the demonstration of validity is not conclusive: the relationship may have been generated by something else about the interview than the variable the research worker thought he was measuring (sheer verbality, extraversion, etc.), but the interview must be measuring *something* reliably to make correlations possible, and since this something is correlated with other variables in a predictable direction, the chances are good that the something being measured is the same thing (or a close cousin) as that which the research worker started out to measure. Perhaps a few examples will make this approach to validity clearer.

In an interview with parents of young children, Sears *et al.* (1953) inquired into the amount and kind of punishment employed. They predicted that the amount of punishment should be related to the amount of aggression children display in fantasy (specifically, in controlled doll-play sessions) and in their play with server, if he knows he is being observed. An example of this is the anecdote told by a candid mother, who with her child had recently been a subject in a behavioral observation study. The two had been observed from behind one-way screens while the child was solving a difficult problem. Several days later, at home, the mother became angry with the child, and the child said: "Mother, why aren't you nice to me at home the way you are over at that playroom?" It is evident that behavioral observation is not intrinsically more valid than the interview as a measure of what would occur in real life, although it may be so in some instances.

What are some of the considerations which bear upon the choice between the two methods? First of all, some behavior, because of conventions about privacy or other reasons, is hardly accessible to observation, so that the interview is the only source of information available. This is true of many aspects of family life, although astonishingly complete observations are sometimes made (see *One Boy's Day,* Barker and Wright, 1951). Perhaps a more difficult problem for behavioral observation is that of sampling the behavior which is the subject of study. The observer can be present only for a limited time and in a restricted range of situations, and the question always arises as to how representative the behavior at these times and in these situations is of the subject's general behavior. Some events occur very rarely in the respondent's life (perhaps only once, as in the case of the death of a parent), and studying reactions to such events through behavioral observation is obviously difficult from the standpoint of timing the observations. In an interview, the interviewer can ask about past events, expectations of future events, and about a wide variety of situations, and the respondent can report what he has done or would do in these circumstances. In this way, the research worker can cover a larger sample of behavior than would be possible with observation, and can learn something fairly quickly and easily about how the behavior he is studying varies with changes in the individual's external situation (always assuming that the respondent can report accurately).

An important question in weighing the relative merits of behavioral observation and the interview is this: to what extent are prestige and status involved in the respondent's activities? The more these are involved, the more distortion one might expect in the interview. (See the discussion of validity above.) Behavioral observation, while subject to the same problem, can sometimes be more free of it, if the observation process is inconspicuous and as much as possible of the real-life situation is present. How aware is the subject of his own behavior and motives? It is in the spheres of behavior where repression and other defense mechanisms are at work that it becomes most important to check interview findings against overt behavior. A field worker reported the following event: A mother was being interviewed concerning her child-rearing practices. During the interview she held her small son on her lap. The child began to play with his genitals. The mother, without looking directly at the child, moved his hand away and held it securely for a while, meanwhile talking to the interviewer. Later in the interview, the mother was asked what she ordinarily did when her child played with himself. She replied that he never did this — he was a very "good" boy. She was evidently entirely unconscious of what had transpired in the very presence of the interviewer. Similarly, people may be unconscious of the ways in which they express aggression, or the ways in which they attempt to get approval from others. When interviewed about their reactions to a moving picture, people may not be able to describe the extent of their emotional arousal, yet managers of large theaters know that for particularly torrid moving pictures they must adjust the air-conditioning apparatus to handle the higher temperatures generated in the theater by the simultaneous emotional reactions of masses of people. It is evident that for such areas of research, interviewing alone is inadequate, although much can be done by using diagnostic questions whose import the respondent does not recognize. On the other hand, the research worker may be working with subject matter in which the verbal expressions themselves are the focus of interest, and in which there is no other behavioral criterion to which direct observation could be applied. For example, in studying ideologies, some aspects of concept formation, or rationalizations, the research worker will necessarily employ some form of the interview as his research tool.

In summary, it appears to the writers that the interview is currently being used uncritically in researches where other techniques would be equally, or more, efficient. On the other hand, some of the most impressive contributions to social science knowledge have been made by studies which employed the interview as their central technique. A growing body of methodological findings is enabling the discriminating research worker to know when to use the interview and how to adapt it to his own research needs.

References

Adorno, T. W., Frenkel-Brunswik, Elsa, Levinson, D. J., & Sanford, R. N. *The authoritarian personality.* New York: Harper, 1950.

Ammons, Carol H., & Ammons, R. B. Research and clinical applications of the doll-play interview. *J. Psychol.,* 1952, *21,* 85-90.

Back, K., Festinger, L., Hymovitch, B., Kelley, H., Schachter, S., & Thibaut, J. The methodology of studying rumor transmission. *Hum. Relat.,* 1950, *3,* 307-312.

Barker, R. G., & Wright, H. F. *One Boy's Day.* New York: Harper, 1951.

Bartlett, F. C. *Remembering.* Cambridge, Eng.: Cambridge University Press, 1932.

Bauer, Alice H. Interviewing Soviet refugees. Unpublished manuscript, Harvard Univ., 1953.

Bingham, W. V., & Moore, B. V. *How to interview.* Rev. Ed., New York: Harper, 1941.

Campbell, A. A. Measuring public attitudes. *J. soc. Issues,* 1946, *2,* 1-69.

Campbell, D. T. The indirect assessment of social attitudes. *Psychol. Bull.,* 1950, *47,* 15-38.

Cannell, C. F. Personal communication, 1953.

Cantril, H. Experiments in the wording of questions. *Publ. Opin. Quart.,* 1940, *4,* 330-332.

Cantril, H. *Gauging public opinion.* Princeton, N. J.: Princeton Univ. Press, 1944.

Chasdi, Eleanor H. Child training among the Zeepi, with special reference to the internalization of moral values. Unpublished doctoral dissertation, Harvard Univ., 1952.

Conn, J. H. The child speaks to the psychiatrist; an introduction to the method of the play-interview. *Occup. Therapy Rehabil.,* 1938, *17,* 231-244.

Covner, B. J. Studies in phonographic recording of verbal material. IV. Written reports of interviews. *J. appl. Psychol.,* 1944, *28,* 89-98.

Crespi, L. P. The cheating problem in polling. *Publ. Opin. Quart.,* 1945-6, *9,* 431-445.

Crutchfield, R., & Gordon, D. H. Variations in respondents' interpretations of an opinion-poll question. *Int. J. Opin. Attitude Res.,* 1947, *1,* 22-31.

Elinson, J., & Cisin, I. H. Detection of interviewer cheating through scales. *Publ. Opin. Quart.,* 1948, *12,* 325.

England, L. R. "Little Kinsey," an outline of sex attitudes in Britain. *Publ. Opin. Quart.,* 1949, *13,* 587-600.

Feldman, J. J., Hyman, H., & Hart, C. W. Interviewer effects on the quality of survey data. *Publ. Opin. Quart.,* 1951-2, *15,* 734-761.

Gardiner, B. B. The meaning and significance of the oil industry to the American people. Unpublished manuscript, 1952.

Getzels, J. W. The assessment of personality and prejudice by the method of paired direct and projective questions. Unpublished doctoral dissertation, Harvard University, 1951.

Goodenough, Florence L. *Anger in young children.* Minneapolis: Univ. of Minnesota Press, 1931.

Gross, N., & Mason, W. Some methodological problems of eight hour interviews. Unpublished manuscript.

Guest, L. A study of interviewer competence. *Int. J. Opin. Attitude Res.,* 1947, *1,* 17-30.

Guetzkow, H. Unitizing and categorizing problems in coding qualitative material. *J. clin. Psychol.,* 1950, *6,* 47-58.

Hammond, K. R. Measuring attitudes by error-choice, and indirect method. *J. abnorm. soc. Psychol.,* 1948, *43,* 38-48.

Heneman, H. G., Jr., & Paterson, D. G. Refusal rates and interviewer quality. *Int. J. Opin. Attitude Res.,* 1948, *3,* 392-398.

Hilgard, E. R., & Payne, S. L. Those not at home—a riddle for the pollsters. *Publ. Opin. Quart.,* 1944, *8,* 245-261.

Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. *Experiments in mass communications.* Princeton, N. J.: Princeton Univ. Press, 1949.

Hyman, H. Do they tell the truth? *Publ. Opin. Quart.,* 1944, *8,* 557-559.

Hyman, H. Problems in the collection of opinion research. *Amer. J. Sociol.,* 1950, *55,* 362-370.

Hyman, H. Interviewing as scientific procedure. In D. Lerner & H. D. Lasswell (Eds.), *The policy sciences.* Palo Alto: Stanford Univ. Press, 1951. Pp. 203-216.

Jahoda, Marie, Deutsch, M., & Cook, S. W. (Eds.) *Research methods in social relations, Vols. 1 and 2.* New York: The Dryden Press, 1951.

Kahn, R. L. The Comparison of two methods of collecting data for social research: The fixed-alternative questionnaire and the open-ended interview. Unpublished doctoral dissertation, University of Michigan, 1952.

Katona, George *Psychological analysis of economic behavior.* New York: McGraw-Hill, 1951.

Katz, D. Do interviewers bias poll results? *Publ. Opin. Quart.,* 1942, *6,* 248-268.

Katz, D., Maccoby, N., & Morse, Nancy C. *Productivity, supervision and employee morale in an office situation, Part I.* Ann Arbor, Mich.: Univ. of Michigan Press, 1950.

Kinsey, A. C., Pomeroy, W. P., & Martin, C. E. *Sexual behavior in the human male.* Philadelphia: Saunders, 1948.

Kluckhohn, C. The use of personal documents in anthropology. In L. Gottschalk, C. Kluckhohn, & R. Angell (Eds.), The use of personal documents in history, anthropology and sociology. *Soc. Sci. Res. Council Bull.,* 1945, *53,* 79-173.

KLUCKHOHN, FLORENCE, STRODTBECK, F. L., ROBERTS, J. M., & ROMNEY, K. A study of cultural orientations. Unpublished manuscript, 1953.

Laboratory of Human Development Field Manual (in preparation). Harvard University, Cambridge, Mass.

LA PIERE, R. T. Attitudes versus actions. *Soc. Forces,* 1934, *13,* 230-237.

LAZARSFELD, P. F. The art of asking why. *Nat'l Marketing Rev.,* 1935, *1,* 26-38.

LAZARSFELD, P. F. The controversy over detailed interviews—an offer for negotiation. *Publ. Opin. Quart.,* 1944, *8,* 38-60.

LEVY, D. M. Use of play technique as experimental procedure. *Amer. J. Orthopsychiat.,* 1933, *3,* 266-277.

LUDEKE, H. C., & INGLIS, R. A. A technique for validating interviewing methods in reader research. *Sociometry,* 1942, *5,* 109-122.

MCNEMAR, Q. Opinion-attitude methodology. *Psychol. Bull.,* 1946, *43,* 289-374.

MEAD, MARGARET Native languages as field-work tools. *Amer. Anthropologist,* 1939, *41,* 189-206.

MEAD, MARGARET Research on primitive children. In L. Carmichael (Ed.), *Manual of child psychology.* New York: Wiley, 1946. Pp. 667-706.

MERTON, R. K., & KENDALL, PATRICIA L. The focused interview. *Amer. J. Sociol.,* 1946, *51,* 541-557.

METZNER, HELEN, & MANN, F. A limited comparison of two methods of data collection: the fixed alternative questionnaire and the open-ended interview. *Amer. sociol. Rev.,* 1952, *17,* 486-491.

MILLER, S. M. The participant observer and overrapport. *Amer. sociol. Rev.,* 1952, *17,* 97-99.

MURRAY, H. A., & MORGAN, CHRISTIANA D. A clinical study of sentiments. *Genet. Psychol. Monogr.,* 1945, *I,* 1-154.

MURRAY, H. A., & MORGAN, CHRISTINA D. A clinical study of sentiments. *Genet. Psychol. Monogr.,* 1945, *II,* 155-309.

NEWCOMB, T. M. The influence of attitude climate upon some determinants of information. *J. abnorm. soc. Psychol.,* 1946, *41,* 291-302.

NORC Questionnaire form from unpublished study, Survey 272, May 1950.

OSS Assessment Staff. *The assessment of men.* New York: Rinehart, 1948.

PARRY, H. J., & CROSSEY, HELEN M. Validity of responses to survey questions, *Publ. Opin. Quart.,* 1950, *14,* 61-80.

PARTEN, MILDRED B. *Surveys, polls and samples.* New York: Harper, 1950.

PAUL, B. D. Interview techniques and field relationships. In A. L. Kroeber (Ed.), *Anthropology today.* Chicago: Univ. of Chicago Press, 1953. Pp. 430-451.

PAYNE, S. L. Interviewer memory faults. *Publ. Opin. Quart.,* 1949, *13,* 684-685.

PAYNE, S. L. *The art of asking questions.* Princeton, N. J.: Princeton Univ. Press, 1951.

PIAGET, JEAN *Introduction to the child's conception of the world.* New York: Harcourt Brace, 1929.

RADKE, MARIAN J. *The relation of parental authority to children's behavior and attitudes.* Minneapolis: Univ. of Minnesota Press, 1946.

RICE, S. A. Contagious bias in the interview: a methodological note. *Amer. J. Sociol.,* 1929, *35,* 420-423.

RIESMAN, D., & GLASER, N. The meaning of opinion. *Publ. Opin. Quart.,* 1948, *12,* 633-648.

ROETHLISBERGER, F. J., & DICKSON, W. J. *Management and the worker.* Cambridge, Mass.: Harvard Univ. Press, 1939.

ROGERS, C. R. The non-directive method as a technique for social research. *Amer. J. Sociol.,* 1945, *50,* 279-283.

SANFORD, F. H. The use of a projective device in attitude surveying. *Publ. Opin. Quart.,* 1950, *14,* 697-709.

SANFORD, F. H., & ROSENSTOCK, J. Projective techniques on the doorstep. *J. abnorm. soc. Psychol.,* 1952, *47,* 3-16.

SEARS, R. R., WHITING, J. W. M., NOWLIS, V., & SEARS, PAULINE S. Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monog.,* 1953, *47,* 135-236.

SHAPIRO, J., & EBERHART, J. C. Interviewer differences in an intensive interview survey. *Int. J. Opin. Attitude Research,* 1947, *1,* 1-17.

SHEATSLEY, P. B. Public relations of the polls. *Int. J. Opin. Attitude Res.,* 1949, *2,* 453-468.

SHEATSLEY, P. B. An analysis of interviewer characteristics and their relationship to performance. I. *Int. J. Opin. Attitude Res.,* 1950, *4,* 473-498.

SHEATSLEY, P. B. An analysis of interviewer characteristics and their relationship to performance. II. *Int. J. Opin. Attitude Res.,* 1951a, *5,* 79-94.

SHEATSLEY, P. B. An analysis of interviewer characteristics and their relationship to performance. III. *Int. J. Opin. Attitude Res.,* 1951b, *5,* 191-220.

SHEATSLEY, P. B. The art of interviewing and a guide to interviewer selection and training. In Marie Jahoda, M. Deutsch, & S. W. Cook (Eds.), *Research methods in social relations, Vol. 2.* New York: Dryden Press, 1951c. Pp. 463-492.

SMITH, H. L., & HYMAN, H. The biasing effect of interviewer expectations on survey results. *Publ. Opin. Quart.,* 1950, *14,* 491-506.

Social Science Research Council Report of Committee on Pre-election forecasts. New York: Social Science Research Council, 1948.

STEMBER, H., & HYMAN, H. Influence of sub-questions on interviewer performance. *Publ. Opin. Quart.,* 1949a, *13,* 310-313.

STEMBER, H., & HYMAN, H. How interviewer effects operate through question form. *Int. J. Opin. Attitude Res.,* 1949b, *3,* 493-512.

STEMBER, H., & HYMAN, H. Interviewer effects in the classification of responses. *Publ. Opin. Quart.,* 1949c, *13,* 669-682.

STEWART, NAOMI, & FLOWERMAN, S. H. An investigation of two different methods for evaluation of interviewer job performance. Paper read at AAPOR, 1949.

STOUFFER, S. A. Some observations on study design. *Amer. J. Sociol.,* 1950, *55,* 355-361.

STOUFFER, S. A., *et al. The American soldier.* Vol. 2. *Combat and its aftermath.* Princeton, N. J.: Princeton Univer. Press, 1949.

STOUFFER, S. A., *et al. Measurement and Prediction.* Princeton, N. J.: Princeton Univer. Press, 1950.

Survey Research Center Introductory statement to coders. Ann Arbor, Mich.: Univ. of Michigan, 1947, ms.

SYMONDS, P. M., & DIETRICH, D. H. The effect of variations in the time interval between an interview and its recording. *J. abnorm. soc. Psychol.* 1941, *36,* 593-598.

U. S. Dept. of Agriculture, Division of Program Surveys. Interviewers' use of reporting forms. Unpublished manuscript, 1945.

VICARY, J. M. Word association and opinion research. *Publ. Opin. Quart.,* 1948, *12,* 81-98.

WESCHLER, I. R., & BERNBERG, R. E. Indirect methods of attitude measurement. *Int. J. Opin. Attitude Res.,* 1950, *4,* 209-228.

WOODWARD, J. L., & DELOTT, J. Field coding versus office coding. *Publ. Opin. Quart.,* 1952, *16,* 432-436.

WYATT, D. F., & CAMPBELL, D. T. A study of interviewer bias as related to interviewers' expectations and own opinions. *Int. J. Opin. Attitude Res.,* 1950, *4,* 77-83.

CHAPTER 13

Content Analysis*

BERNARD BERELSON
Ford Foundation

In the communication process a central position is occupied by the content. By communication content is meant that body of meanings through symbols (verbal, musical, pictorial, plastic, gestural) which makes up the communication itself. In the classic sentence identifying the process of communication — *"who* says *what* to *whom, how,* with *what effect"* — communication content is the *what.* This chapter deals with the analysis of what is said in communication of various kinds.

Since the content represents the means through which one person or group communicates with another, it is important for communication research that it be described with accuracy and interpreted with insight. Communication content is so rich with human experience, and its causes and effects so varied, that no single system of substantive categories can be devised to describe it. However, a scientific *method* has been developed — and is being developed further — for describing various facets of communication content in summary fashion. That method is called *content analysis.* It has been used to answer such diverse questions as these:

How have the slogans of May Day propaganda in the U.S.S.R. changed during the Soviet regime?

How is the writer's personality structure reflected in what he writes?

How do the values in American plays differ from those in German plays of the same time?

What is the pattern of movement in the deliberations of a small group? What is the pattern in social case work?

How are minority ethnic groups treated in short stories in popular magazines?

How can communications suspected of subversion be tested for their "propaganda" component?

This report on the technique of content analysis attempts to survey critically the applications of content analysis, to relate it to marginal developments in other fields, to note the major types and categories of analysis, and to review certain technical problems. After a general introductory section, the report discusses the major uses of content analysis, the units and the categories of analysis, and such technical matters as sampling, reliability, and modes of inference.

DEFINITION OF CONTENT ANALYSIS

What is meant by the term "content analysis"? Review of several definitions which have appeared in the technical literature will serve to identify its major characteristics:

> Systematic content analysis attempts to refine more casual descriptions of the content, so as to show objectively the nature and relative strength of the stimuli applied to the reader or listener. (Waples and Berelson, 1941, p. 2)
>
> A social science sentence may be called one of "content analysis" if it satisfies all of the following requirements: (1) it must refer either to syntactic characteristics of symbols . . . or to semantic characteristics . . . (2) it must indicate frequencies of occurrence of such characteristics with a high degree of precision. One could perhaps define more narrowly: it must assign numerical values to such frequencies. (3) It must refer to these characteristics by terms which are general . . . (4) it must refer to these characteristics by terms which occur . . . in universal propositions of social science. One may consider adding to this definition another requirement: (5) a high precision of the terms used to refer to the symbol characteristics studied. (Leites and Pool, 1942, pp. 1-2)
>
> The content analyst aims at a quantitative classification of a given body of content, in terms of a system of categories devised to yield data relevant to specific hypotheses concerning that content. (Kaplan and Goldsen, 1943, p. 1)
>
> "Content analysis" may be defined as referring to any technique for the *classification* of *sign-vehicles;* which relies solely upon the *judgments* —which, theoretically, may range from perceptual discriminations to sheer guesses—of an analyst or group of analysts as to which sign-vehicles fall into

---

*At the request of the editor, Dr. Berelson has prepared a condensation, by half, of his *Content Analysis in Communication Research,* The Free Press, 1952. Few changes have been made in the process of condensation.

> which categories; on the basis of *explicitly formulated rules;* provided that the analyst's judgments are regarded as the report of a *scientific observer.* The results of a content analysis state the frequency of occurrence of signs—or groups of signs —for each category in a classification scheme. (Janis, 1943, p. 429)
>
> . . . The technique known as content analysis . . . attempts to characterize the meanings in a given body of discourse in a systematic and quantitative fashion. (Kaplan, 1943, p. 230)

This group of definitions provides six distinguishing characteristics of content analysis:

(1) It applies only to social science generalizations: Leites and Pool.

(2) It applies only, or primarily, to the determination of the effects of communications: Waples and Berelson.

(3) It applies only to the syntactic and semantic dimensions of language: Leites and Pool.

(4) It must be "objective": Waples and Berelson, Leites and Pool, Janis, Kaplan.

(5) It must be "systematic": Leites and Pool, Kaplan and Goldsen, Kaplan.

(6) It must be quantitative: Waples and Berelson, Leites and Pool, Kaplan and Goldsen, Janis, Kaplan.

As we shall see, the first and second of these characteristics define the field of content analysis too narrowly. The review of the literature will show that it has been applied successfully in other fields than the social sciences and for other purposes than the description of the effects of communications upon readers and listeners. But the other four characteristics are required for a proper definition of content analysis.

The *syntactic-and-semantic requirement* is meant to rule out the analysis of communication content for the pragmatic dimension of language (the third branch of semiotics, the general science of signs, as developed by Charles Morris). That is, content analysis is ordinarily limited to the manifest content of the communication and is not normally done directly in terms of the latent intentions which the content may express or the latent responses which it may elicit. The results of content analysis, however, frequently serve as the basis for these "interpretations" of the latent content.

The *requirement of objectivity* stipulates that the categories of analysis should be defined so precisely that different analysts can apply them to the same body of content and secure the same results. Like the first requirement, this ordinarily limits content analysis to the manifest content.

The *requirement of system* contains two different meanings. In the first place, it states that *all* of the relevant content is to be analyzed in terms of *all* the relevant categories, for the problem at hand. This requirement is meant to eliminate partial or biased analyses in which only those elements in the content are selected which fit the analyst's thesis.

The second meaning of "system" is that analyses must be designed to secure data relevant to a scientific problem or hypothesis. The results of a content analysis must have a measure of general application.

The *requirement of quantification,* the single characteristic on which all the definitions agree, is perhaps the most distinctive feature of content analysis. Of primary importance in content analysis is the *extent* to which the analytic categories appear in the content, that is, the relative emphases and omissions. Now this requirement of quantification does not necessarily demand the assignment of *numerical* values to the analytic categories. Sometimes it takes the form of quantitative words like "more" or "always" or "increases" or "often." Although results of this kind may be appropriate for certain studies, it should be recognized that such terms are just as "quantitative" as the terms 37 or 52%; they are only less exact and precise. In most applications of content analysis, numerical frequencies have been assigned to occurrence of the analytic categories.

This review of the distinguishing characteristics of content analysis, then, results in the following definition: *Content analysis is a research technique for the objective, systematic, and quantitative description of the manifest content of communication.*

THE HISTORY OF CONTENT ANALYSIS

The output of content analysis studies has sharply increased in every five-year interval over the past thirty years. From about two studies a year in the early decades of the century, content analysis studies have increased to three a year, then four, eight, twelve, and in the most recent years to an average of about 25 a year. Presumably the field is still expanding.

Content analysis was first used by students of journalism (and later by sociologists) to study the content of American newspapers. This work centered in the School of Journalism at Columbia University, and culminated in Malcolm Willey's study, *The Country Newspaper,* in 1926. Almost exclusively, such early studies employed straight subject-matter categories (e.g., domestic affairs, politics, labor, crime, divorce, sports, etc.).

The other early concentration of content analysis studies occurred in the field of literature, with the analysis of various stylistic features in English poetry and prose (e.g., sentence length, predications, conjunctions, etc.). The contribution of this kind of study was apparently not valued highly by students of literature; it was many years before it was applied again to problems of literary style by Rickert, Spurgeon, Miles, and others.

During the late 1930's content analysis was given an important stimulus through the work of Harold D. Lasswell and his associates, as well as through the growing interest in propaganda and public opinion and the emergence of radio as a great mass medium of communication. Viewing communication problems within the framework of a theory of politics, Lasswell introduced new problems, new procedures, and new categories into the field (and thereby imparted a *political* complexion to it). The concern of Lasswell and others with problems of public opinion and propaganda led naturally to the study of communication and thence to content analysis. At about this time, the rise of radio also directed attention to content analysis, most notably under Paul F. Lazarsfeld of the Bureau of Applied Social Research at Columbia University. Also during this period, the professional interest in adult education promoted several researches into the readability of materials. During World War II, content analysis was employed by several government departments charged with responsibilities in the field of mass communication. The Experimental Division for the Study of Wartime Communications, in the Library of Congress, the Analysis Division of the Foreign Broadcast Intelligence Service (Federal Communications Commission), and the Organization and Propaganda Analysis Section of the Special War Policies Unit, Department of Justice, all employed content analysis techniques in their investigations of the propaganda output of various countries, organizations, or individuals. In addition, content analysis has been given certain commercial applications.

THE USES OF CONTENT ANALYSIS

The method of content analysis has been applied to so large and diverse a group of materials, with respect to so large and diverse a set of problems, that it is not easy to order the uses in a single classification. However, seventeen types of uses (applications, functions) of content analysis have been identified.

This set of uses is presented here as a list, not as a classification system. That is, there is no claim that the uses form a logically coherent organization or that they are classified on a single dimensional base. This classification has been induced from the major purposes of the studies and as such it presumably reflects the central problems to which the method has been applied. If the list is not sufficiently coherent and systematic, neither is the field.

The uses of content analysis are ordered under major headings which progressively refer to the *characteristics* of communication content, the *causes* of content, and the *consequences* of content.

CHARACTERISTICS OF CONTENT: SUBSTANCE

To describe trends in communication content. "The most valuable use of studies of content . . . is in noting trends and changes in content. Systems of classification may be inadequate and unstandardized; nevertheless, if a system is used consistently over a time period valuable facts may appear" (Albig, 1938, p. 349). This quotation indicates the importance sometimes attached to this application of content analysis, although it seems an overstatement in view of the more advanced uses of content analysis. The classification into a single set of categories of similar samples of communication content taken at different times provides a concise description of content trends, in terms of relative frequencies of occurrence. Such descriptions of trends are often useful in themselves; in addition, they provide data which can be correlated with corresponding changes on the part of the communicator or the audience.

The obvious importance of this application of content analysis has led to many such studies. The contributions can be illustrated by a few examples:

In thirty Minnesota weekly papers, the percentage of shop-set news originating in the locality and trade area rose from 28% in 1860 to over 80% in 1890 and thereafter. Personal news increased from less than 6% in 1860 to almost 40% in all years after 1900 (Taeuber, 1932).

There was relatively little difference in the extent to which economic, nationalistic, militaristic, or internationalistic motivations were employed in the McGuffey readers and in a sample of school readers in 1930, but there was a considerable difference in the relative use of religious motivations (Estensen, 1946).

The propaganda content of May Day slogans in the Soviet Union from 1924 to 1943 moved from "universal revolutionary" symbols to "national" symbols. (Yakobson and Lasswell, 1949).

Whereas only about half the articles on child development in the *Ladies' Home Journal* and *Good Housekeeping* in 1904 and 1910 were based on "scientific authority" rather than "personal opinion," the proportion in 1940 had risen to fully 98% (Ojemann, 1948).

As compared with World War I, propaganda in World War II was less emotional, less moralistic, and more truthful (Kris and Leites, 1947).

Such trend studies provide a valuable historical perspective against which the current content of the communication media can be more fully understood. Indeed, some students have suggested the organization of continuing trend studies of the major media on various topics in order to provide a background against which particular studies could be interpreted. In this connection, it is important to point out that most of the existing trend studies have dealt only with subject-matter categories, although the addition of other categories (e.g., direction, authority, emotionality) would make such studies even more valuable.

To trace the development of scholarship. The preceding section has discussed content trends in the general media; this section deals with the special case of content trends as indices to the development of scholarly interests and activities. Content analyses of the literature in various scholarly and scientific fields have been used to describe their development. An analysis of the development of physics in England, Germany, and France during the 18th and 19th centuries produced certain "wave-like fluctuations of creative productivity," decades in duration (Rainoff, 1929). The changing nature of botanical research from 1890-1930 is revealed in the study of the content of a group of American journals in that field (Stevens, 1932). Similarly, the development of sociological interests in the United States over the past fifty years is indicated by the distribution of articles in the *American Journal of Sociology* (Shanas, 1945). In recent years the *Journal* has increased its attention to problems of population, races and nationalities, and methods of research at the expense of papers dealing with the theory and history of sociology and with social reform. Such analysis not only describes the developmental trends of a discipline, but also supplies the definition or composition of the field at any particular time. A similar analysis of psychological journals was carried out by Allport and Bruner (1940).

The particular assumptions involved in this use of content analysis must be recognized and validated. Most such studies rely upon a single source of data, or a small group of sources, over a relatively long period of time, and the continued adequacy and consistency of the data may require special attention. For example, a few general journals may not contain a valid sample of the research literature of a given field over a long time because of the later addition of more specialized journals representing branches of the field. Or, again, changes in the content of a relatively few journals in a field may reflect changes in editors or editorial policies rather than genuine, broadly recognized shifts of interest within the field itself. In some cases, results may be distorted by the national origin of the source materials, as, for example, in analyses of "great men" by means of space measures in the *Encyclopaedia Britannica* and/or the *Americana*. And trend studies may be distorted by the tendency of sources to give more attention to the work of the period immediately preceding their own. Even changes in classification policy may be a disturbing influence; for example, articles on the borderlines between physics and mathematics or between sociology and political science may have been assigned to different fields at different times. However, precautions can be taken against most of these difficulties, and then content analysis becomes a useful tool for describing the composition and development of scholarly and scientific fields.

To disclose international differences in communication content. Systematic analysis can disclose differences in the communications available in various countries. What news of common interest and relevance is carried in newspapers or on the air, and how is it treated? How different are international versions of the same event, such as wars or international crises in which nations have equal investments? What are the central differences between, say, American, British, French, and Russian films? Such questions can be answered through comparative content analyses which will reveal differences in the focus of attention of national audiences.

An intensive analysis (though not particularly systematic or rigorous) revealed the differing treatment of this country's wars in American history textbooks and in the textbooks of the former enemy countries (Walworth, 1938). The versions of the Revolutionary War in American and British texts or of World War I in American and German texts are compared in order to reveal the common or divergent views of history presented to the youth of different countries. The relevance of such presentations for international attitudes is clear. So is the record: in virtually every case, the divergencies are predominant. Take, for example, the American and British versions of the Revolutionary War. In simplified form, the situation is described thus (Walworth, 1938, pp. 3-20):

American	*British*
King George III is a blundering bigot, obstinate and narrow-minded.	King George was a man of moderate ability with a firm will, real courage, and a high sense of royal obligation to the public.
America had contributed to British prosperity.	The colonies were a liability rather than an asset to the British Empire.
The "Intolerable Acts" passed by the British Parliament were the beginning of hostilities.	Acts of violence committed by the radical element in the colonies made military action necessary.
The Colonies were fighting the cause of traditional English liberty by their action.	(omitted)
(omitted)	King George made a last attempt at reconciliation.
The American defeat at Bunker Hill was a moral victory, for it showed that the raw Colonial troops could face the regulars without flinching.	. . . The small English army suffered as heavy losses with as steady courage as has ever been shown on prouder fields of honour.

Although this particular study is based upon comparative quotations culled from selected works, with only a few rough references to central tendencies, there is no question but that systematic techniques of analysis can be applied to this problem.

Two unusual studies of German-American differences come to similar conclusions, although they involve widely differing materials. In the one case, the literature of the Hitler Youth and that of the Boy Scouts of America were analyzed in terms of goals and justifications (Lewin, 1947). There was no difference in emphasis upon such ends as physical fitness, discipline, or achievement, but there was more stress in the German material on national loyalty and identification and less on altruism and creativity. In addition there was considerable difference in the bases for recommending desirable ends to the members of the two organizations. The Hitler Youth recommendations were justified primarily on the basis of membership in the national community (66%), while the Boy Scouts justified ends in terms of the individual's own satisfaction (47%).

In the other study, essentially the same thing was done for the most popular dramas in the two countries (McGranahan and Wayne, 1948). Whereas the major themes in American plays were love (60%) and morality (36%), in German plays they were ideals (44%) and power (33%). In American drama the "level of action" was almost exclusively personal (96%); in German drama it was evenly divided between ideological and personal affairs (51% and 47%). Now while both these studies can, of course, be used to illustrate and perhaps document differences in national character, they also can be taken as descriptions of what was presented to similar groups in the two countries.

To compare media or "levels" of communication. Different media and different "levels" of communication not only attract different audiences but they also treat the same topics in different ways. For both of these reasons, the comparative analysis of communication content is among the most promising applications of content analysis, and a number of interesting studies have appeared in recent years.

For example, from a detailed study of 24 "good" novels and the films adapted from them are developed the major types of changes found in such adaptations — with respect to dialogue, character, action, scene, setting, sequence — and the (inferred) reasons for them (Asheim, 1949). Simply the amounts of attention given to the central plot components in book and film serve as an index to the fidelity of the transcription from one medium to the other:

MOST FAITHFUL

Pride and Prejudice

	Book	*Film*
The Elizabeth-Darcy romance	48%	47%
The Wickham Story	23	14
The Jane-Bingley romance	13	12
The story of Mr. and Mrs. Bennett	12	20
The character study of Collins	12	14
Minor characters	9	9
The character study of Lady Catherine	6	13
Pure description	5	4

LEAST FAITHFUL

The House of the Seven Gables

	Book	*Film*
The story of Hepzibah and Clifford as brother and sister	30%	0%
The story of Hepzibah and Clifford as lovers	0	43
The story of Phoebe and Holgrave	28	21
The story of Judge Pyncheon's plot against Clifford	19	0
The story of Judge Pyncheon and the counterplot of Clifford	0	35
Commentary	16	5
General descriptions	12	14
Lesser characters	7	8

In addition, the study disclosed the probable bases for various content changes which were made in many adaptations. Thus, for example, the technical requirements of film-making produced an "impersonal omniscient" viewpoint in 23 of the films, whereas it had occurred in only nine or ten of the books.

Another study is distinctive for its comparison of different "quality levels" dealing with American history; it investigated the extent to which the basic findings of research monographs were incorporated in secondary school texts on American history (Blythe, 1932). About thirty new viewpoints, emphases, and discoveries in historical scholarship from 1893 to 1928 were checked, e.g., Turner's interpretation of the significance of the frontier, Mahan's emphasis on sea power in the War of 1812, Beard's economic interpretation of the Constitution, Becker's view of the American Revolution as a double revolution, and so on. The typical finding was incorporated in only one-third of fifteen then-recent texts, and individual texts included only from five to nineteen of the findings. On the average, it took about six years for a development from historical research to find its way into history textbooks for secondary schools. Because different groups of people read and listen to different kinds of communications on the same subject, the comparative application has considerable potential value as a method for describing the differences in content to which they are exposed. In addition, it is particularly useful in comparing different "levels" of communication content.

To audit communication content against objectives. Every communication outlet has an objective or a set of objectives, whether implicit or explicit. *One* measure of the quality of the content is the extent to which it faithfully expresses such objectives.

Not many studies have been done primarily with reference to this function of content analysis. A large number have compared content emphases with the standards of the *analyst* (to be discussed in the next section), but only a small number have checked content against the communicator's *own* objectives. Sometimes such analyses are made by "outsiders" on the basis of the *assumed* objectives of the communicators. Thus an analysis of the "educational" content of television was carried out in early 1951 by a group interested in promoting such programs. On the assumption that the television industry would argue, in F.C.C. hearings, that it was already fulfilling this purpose adequately, the total New York output for one week was analyzed for type of program and the findings became an impressive exhibit (Smythe and Horton, 1951). For example, public issues received 2% of the total program time, public events 1%, and information programs 3%, as compared with 10% for crime drama and 14% for variety.

Content analysis can be employed to correct the constant, often inadvertent, and perhaps inevitable omissions from and mis-emphases in communication content, in terms of the objectives and directives of the communicator.

To construct and apply communication standards. Content analysis can describe communications but it cannot, *per se,* evaluate them. Evaluation necessitates the acceptance of a standard, or standards, with which the communication content is then compared by means of content analysis.

Content analysis has been used in three methods of evaluating communications: (1) evaluation of performance against such *a priori* standards as "balance" or "social purpose"; (2) evaluation of performance by comparing one body of content with another; and (3) evaluation of performance by comparing content with a noncontent source. In each case, the standards are advanced by someone other than the communicator himself — which distinction differentiates this use from the audit function of content analysis.

Perhaps the major recent application of *a priori* standards to communication content was made for the Royal Commission on the Press (1949). This British group utilized a content analysis, of subject and form mainly, ". . . in order to ascertain how great was the extent of suppression, distortion, and inaccuracy . . ." (p. 107). Partly on the basis of this research, reported in several tabulations in the appendix to the report, the Commission concluded ". . . that the political factor in the selection and presentation of news is apparent in all the national papers . . ." (p. 109). Other examples have dealt with the treatment of labor in radio news programs (Sussman, 1945), the treatment of the Taft-Hartley Act in 50 periodicals (Ash, 1948), and the incidence of local news in county weeklies (Willey, 1926).

The second basis of evaluation involves the construction of standards internally, i.e., within the content itself, in which case comparisons are made between one body of content and another. For example, in a study of "ethical and unethical behavior by representative newspapers," the performances of several papers were evaluated against the average "socialization-sensationalism" profile of a group of papers highest in "socialized news attention" (Kingsbury and Hart, 1937). The profiles were secured by the analysis of front-page headline space devoted to three "highly social" subject matters (foreign problems, U.S. international, and citizenship), and three "extremely sensational" ones (money-sensational, sex, and money-sex). By the subtraction of the latter from the former, a single "socialization-sensationalism score" was secured for each of some forty newspapers ranging from the *United States Daily* (32.8) and the *Christian Science Monitor* (30.3) to the *New York Daily News* (—14.2) and the *New York Daily Mirror* (—17.3). Thus, the papers were compared on this standard with one another as well as with the norm of the "better" newspapers.

Finally, the third method of evaluation — against an outside noncontent standard — was used in assessing the frequency with which members of different ethnic groups appeared in magazine stories (Berelson and Salter, 1946). In this case, the distribution of the fictional characters was compared with the distribution of the actual population. "Census data only accentuates the differential treatment accorded 'natives' and 'minorities' in the stories . . . Every 'minority' group appears less frequently in the stories than in the country. Only 'The Americans' appear more frequently" (p. 173).

Another such comparison, useful in evaluating newspaper performance, was employed to determine the extent to which British papers reported the major facts on an important event, the first year's progress of the National Coal Board (Royal Commission, 1949). The analysts constructed a master list of 33 major facts — "what actually happened" — against which they checked the coverage of the papers. The scores for nine papers ranged from 23½ for the *Times* to 3 for the *Daily Worker.*

The discussion of standards for mass communications will almost certainly increase in coming years, and in this discussion the use of content analysis can be extremely valuable in evaluating the media. Once the nature of the standards is clearly and explicitly formulated, content analysis can be usefully applied to measure the extent to which the specified standards are met. Potentially, at least, content analysis can provide objective data on performance to compare with the agreed-upon norms stated in the form of communication standards.

To aid in technical research operations. To some extent, content analysis has been used as one step in a series of research operations. The major use here involves the coding of qualitative materials secured through sample survey or experimental methods.

In recent years there has been considerable development in the methods of collecting various kinds of information, opinions, and interaction data by means of personal interview or recorded observation. One such method is the intensive or open-end interview, in which the respondent is asked a question without being given a checklist or other structured form in which to answer. By means of probes and other devices, the respondent is encouraged to reply at some length. For certain purposes, this type of question is indispensable. However, its use raises the problem of the systematic and objective analysis of the answers. "It can safely be said that the proponents of the open-end interview technique have made much more progress in the conduct of the interviews than in their statistical analysis" (Lazarsfeld, 1944).

Strictly speaking, the quantification aspect of content analysis is seldom applied to such interviews — partly because the coding rarely requires such intensive analysis, partly because the materials would not ordinarily support it (i.e., they are too far from a verbatim account of what the respondent actually said), and partly because the meaning of the quantification would rest upon dubious assumptions about repetition, extension of remarks, etc. The coding procedure is similar to content analysis in that it involves the codification of verbal materials which can take many diverse forms and still "mean" the same thing, in terms of the categories under which they are subsumed. But there the similarity usually ends. Thus coding is based upon the presence or absence of the categorized material, not upon its frequency — upon *whether*, not *how much*.

However, in one form of this material a variant of content analysis has been used to classify the subjects on the basis of their total response to the interview, with consideration of the point in the interview at which certain responses were made. In a study of ethnic prejudices of a group of veterans, the respondents were classified into the intense (anti-Semitic or anti-Negro), the outspoken, the stereotyped, and the tolerant according to whether they made spontaneous restrictive remarks, elicited restrictive remarks, stereotyped responses, or explicit references of a tolerant nature (Bettelheim and Janowitz, 1950). In another case, the category of "attitude toward political parties in general" was coded on the basis of the pros and cons secured in answer to several questions dealing with political matters of various kinds.

More direct application of content analysis as a research aid is found in some recent analyses of group behavior (Chapter 10). A "profile of group interaction" has been developed by the quantitative analysis of statements made by group members, in terms of such categories as these: shows solidarity, shows tension release, agrees, gives suggestion, gives opinion, gives orientation (and their counterparts). Hypotheses on group behavior under varying conditions are tested by means of this use of content analysis (Bales, 1950).

CHARACTERISTICS OF CONTENT: FORM

To expose propaganda techniques. Content analysis is often identified with propaganda analysis. In one sense, a great number of content studies — certainly the majority of them — can be considered propaganda analyses, if by propaganda is meant the deliberate attempt to influence attitudes or behavior on controversial public issues. Even the study of the treatment given education in newspaper editorials becomes an analysis of propaganda, and its categories become "propaganda techniques." In this way, then, almost every category used in content analysis can be considered a "propaganda technique" or a category of "propaganda analysis," since almost every category has been used at one time or another in analyses of material designed to influence attitudes on controversial issues.

As the field is usually identified, however, certain "techniques" can be illustrated from the "propaganda analyses" presented in the literature. The first type of category that appears in many such analyses is the theme or, as it is often called, the "appeal." Now "appeal" is a term which ordinarily refers not to characteristics of the content, but to characteristics of either the intent or the effect of the communication; the corresponding term for content is probably "theme." Various content formulations, or themes, may carry the same "appeal"; similarly, the same theme may carry different "appeals" for different audiences. As already indicated, it is preferable to analyze communications in content categories rather than in terms of presumed intents or effects, and then to make the latter type of interpretation on the basis of analytic results. An illustration of this process is provided by a study of propaganda techniques used in World War I (Lasswell, 1927). The analysis of the themes used by the opposing sides illustrated the four major objectives of wartime propaganda: to mobilize hatred against the enemy, to preserve the friendship of allies, to preserve the friendship and if possible to secure the cooperation of neutrals, and to demoralize the enemy. Within this framework the effectiveness of the propaganda was evaluated: "The British were amazingly successful in the development of humanitarian war aims. The Germans aroused much resentment and suspicion abroad by talking about a war of German Kultur, and by underplaying the humanitarian ideal . . . British public men began to talk about a war to end war long before the German statesmen learned this vocabulary . . . The Germans were never able to efface the initial impression that they were aggressors. This was due in part to the stupidity of their own appeals" (Lasswell, pp. 196-197). A more molecular and quantitative approach to the analysis of political communications is provided by White (1942, 1949).

Another set of categories used in propaganda analysis — perhaps the set most frequently associated with the term — deals with the "tricks of the trade." They are the various special devices of which the subject is not ordinarily aware. Perhaps the most appropriate illustrations of such analysis are found in the work of the Institute for Propaganda Analysis. Seven such devices are described and illustrated in a study of the speeches of Father Coughlin (Lee and Lee, 1939). This list was later expanded to eleven "tricks of the trade":

A. Techniques of basic procedure
  1. Selecting the issue
  2. Case-making or card-stacking
  3. Simplification

B. Use of omnibus words (basic propaganda shorthand)
  4. Name-calling
  5. Glittering generality

C. Techniques of identification
  6. Transfer
  7. Testimonial
  8. Plain folks
  9. Bandwagon

D. Strategic techniques
  10. Hot potato
  11. Stalling

Although such devices may be useful in making people propaganda conscious, it should be clearly recognized that they do not constitute the "whole" area of propaganda analysis. Little has been done to define these categories precisely, and there has been no effort to correlate systematically the appearance or frequencies of these categories with the effectiveness of the material.

The source of assertions is another propaganda characteristic which has been dealt with in content analyses. This category was included in the study of German short-wave propaganda through a comparison of certain prestige devices in British and German propaganda dealing with the same event, i.e., the Altmark incident (Bruner, 1941). As it happens, the two countries used authoritative symbols quite similarly in this case.

This brief review of some categories that have been used in "propaganda analysis" cannot omit the characteristic which is often associated with the term "propaganda," namely, the emotional content of the material. Some attempts have been made to measure systematically the relative emotionalization of various propaganda, usually by counting the frequency of emotional, or sentimentalized, terms. An illustration of this kind of propaganda analysis is provided by a study of the 1940 presidential campaign, which revealed that significantly more emotional content appeared in the material dealing with Roosevelt than in that dealing with Willkie (Waples and Berelson, 1941). The same procedure revealed that the emotional content of newspaper material on the campaign was higher than that of radio or magazine material and that the emotional tone of propaganda during the last month of the campaign was twice as great as that during the relatively quiet period of early August.

The analysis of "propaganda techniques" may take a large number of forms and employ a large number of different categories. It should not be considered as restricted to the emotional component or to the "tricks of the trade." Just as effective propaganda for various audiences may necessitate many and complex appeals, so must content analysis designed to identify the efficient influences within it be varied and complex.

To measure "readability." One of the two main uses of content analysis by educators is the measurement of the "readability" of printed materials; textbook analysis is the other main use. What are the concrete factors that distinguish material that is easy or hard to read and comprehend? What are the stylistic differences, in these terms, between *True Story* and *The Atlantic Monthly?* The answers to such questions are used by educators in the selection and even the production of appropriate books for readers of varying ability.

In a sense, the development of readability measures is a story of increasing differentiation. In the beginning of such analysis the "word" was focused upon. The first studies of readability in the mid-1920's classified materials simply on the basis of the proportions of "easy" and "hard" words they contained, as indicated by a standard word list (usually the Thorndike list). In the following decade additional factors were investigated — sentence length, simple sentences, indeterminate clauses, prepositional phrases. Then followed a thorough investigation of no fewer than 82 assumed elements of difficulty, of which five were finally selected as correlating best with tests of comprehension: different words, different hard words, personal pronouns, prepositional phrases, and sentence length (Gray and Leary, 1935). This set of content characteristics was next reduced to three (hard words, prepositional phrases, sentence length), and they were included in a formula which directly expressed readability in terms of the years of schooling required for easy reading of the passage (Lorge, 1944). At about the same time another widely used formula contained a different set of three factors — sentence length, personal references, and affixes (Flesch, 1943). And another formula based on the number of words not on a certain list (of 3000 words) and on average sentence length was constructed (by Dale and Chall, 1948) to overcome criticisms of Flesch's formula, which was itself later divided into two facets of readability — "reading ease" and "human interest" (Flesch, 1951, pp. 6, 10).

There is an important limitation to the applicability of readability analyses. They have been validated against reading comprehension tests of general readers representing different educational groups. Hence, they apply to the *general* reader with *general* interest, and they do not take into account the effect upon readability of the reader's own special interests. For example, a poorly educated mechanic could comprehend some technical material which would prove too difficult for a well-educated businessman. Readability is a function of the reader's interest and experience, and this may seriously qualify its mechanical application. Furthermore, the formal categories in terms of which readability is normally measured do not take into account such factors as organization of the material, nature of the content, format, etc. (for a general criticism in these terms, see Kearl, 1948).

To discover stylistic features. The problems of literary and oratorical style have attracted students since ancient times, and countless volumes have been written about them. What are the distinctive stylistic features of poetic diction? How do writers differ in literary style? How have rhetorical devices changed over a period of time? When were particular works written, and by whom? Such questions have traditionally been analyzed by students of literature in so-called "qualitative" terms. But from time to time, venturesome persons have experimented with the quantitative analysis of style. Such analysis has had an uneven reception in the field, but on occasion it has made important contributions to stylistic analysis and for some less complicated problems it is more or less standard. Consideration will be given first to its application to language and literature, and then to rhetoric and oratory.

Language and literature. There are four major areas in which content analysis is applied to differentiate stylistic patterns. The first is an *entire language.* Various linguistic studies have been made of different features of a whole mode of speech. For example, there is the analysis of the grammatical structure of American English, in both the standard and vulgar forms (Fries, 1940). Certain recommendations about the teaching of English grammar are based upon the frequencies with which patterns of usage appear in writing and speech.

The second area of application differentiates the *stylistic patterns of various periods.* This has been a central concern of one of the most recent, and most extensive, applications of quantitative analysis to literary style, and specifically to problems of poetic vocabulary (Miles, 1943, 1946, 1951). The ratio of verbs to substantives (nouns, plus adjectives) is taken as the indicator of a significant trend: "fairly steady rise toward an extreme, marked by adjectival quantity, and then a slower decline, moving toward moderation and not far beyond it . . ." (Miles, 1951, p. 520).

The third area involves the differentiation of style by *types of discourse.* Certain stylistic features of types of written discourse have been investigated by means of quantitative analysis. For example, substantive-verb ratios are used to characterize different types of poetry (Miles, 1951). The claim is, of course, that such crude measures are only indicative of more basic features of the poetry.

"Substantival dominance"	*Type*	*Examples*
2 to 1	"Metaphysical"	Wyatt, Donne, Emerson, Millay
3 to 1	"Romantic"	Wordsworth, Poe, Arnold, Yeats
4 to 1	"Classical"	Spenser, Dryden, Crane

The fourth area is the most particular, the *individual author.* Content analysis has been applied to the substantive meanings in literature as well as to its structural elements. Classification of the images in Shakespeare's plays illustrates the wide range of figures upon which he drew and also provides a comparison with his contemporaries (Spurgeon, 1936); in addition, the imagery of particular plays or groups of plays clarifies the way the poet achieved his effect, thus:

Play	*Dominant image*
Early historical plays	growth and deterioration in a garden
Romeo and Juliet	light in a world of darkness
Hamlet	disease and sickness
Macbeth	"ill-fitting garments"
Othello	"animals in action, preying upon one another"
Lear	"human body in anguished movement"

Two special cases of the quantitative analysis of formal stylistic elements, usually in poetry, are reasonably well accepted by literary scholars. One is the problem of disputed authorship and the other that of establishing the correct chronology of an author's works. These problems are attacked through the analysis of such stylistic characteristics as type of rhyme, feminine endings, size of vocabulary, compound words, etc. Thus one scholar assigns the play *Edward III* to Shakespeare on the basis of the relative frequency of compound words (Hart, 1934). Another, on the basis of vocabulary and sentence-length evidence, concludes that "the results . . . almost exclude the possibility of Gerson as the author (of *De Imitatione Christi*), but are entirely consistent with the authorship of Thomas à Kempis" (Yule, 1941, p. 278).

Rhetoric and oratory. Content analysis has been applied less frequently in the study of rhetorical and oratorical style. However, it is unexpectedly found in a study of the appearance in the Attic orators of types of rhetorical arguments from Aristotle's Rhetoric (Palmer, 1934). The number of illustrations of the twenty-eight types ranges widely; for example:

Type of Argument	*Number*
a fortiori	several hundred
the appeal to authority, the result of a decision, consideration of inducements and deterrents, and the juxtaposition of inconsistencies	more than a hundred
the argument from division, argument by induction and resulting deduction, and argument from motives	30 or more
exposure of inconsistencies, the reversal of an earlier choice, and the denial of an interpretation	from 10 to 20
argument by identification and association, and the probability of improbability	less than 10
argument through ambiguity, and play upon names	none

Other applications in this area have been made to political materials. One study of 145 speeches made in presidential campaigns from 1884 to 1920 disclosed that nearly 40% of them contained some class appeal (most frequently to labor), that about 25% of them contained an appeal to party loyalty, and that more time was spent in attacking the opposition than in pointing out the achievements of the speaker's own party (Hayworth, 1930). An intensive analysis of the style of Woodrow Wilson's political speeches, after-dinner addresses, and sermons showed the relative frequency of rhetorical figures, with metaphor, rhetorical questions, and personification appearing most often (Runion, 1936). A more sophisticated and clever analysis of Wilson's style shows how his personality was expressed in such stylistic traits as "aristocratic affectations," "learned addictions," etc. (Hale, 1920).

PRODUCERS OF CONTENT

To identify the intentions and other characteristics of the communicators. This has been a major use of content analysis. The basic logic of this application of content analysis is this: the content has such-and-such characteristics, therefore the communicators have such-and-such intentions. As in the case of cultural patterns, the relevant characteristics of the communicators are studied indirectly by this method only because the content is more readily available.

Analysis of correspondence in content may support inferences about collaboration in intent. During the war, official government analyses were made in order to reveal the extent to which the German and Italian propaganda ministries were cooperating, if at all (Berelson and de Grazia, 1947). The correspondence between the Berlin and Rome radios was measured in several ways — by the amount of attention given to particular events, by the broad allocation of radio time, by the treatment of news and comment, by the similarity and the adoption of propaganda themes, and by identifications of the self. In every case, inferences about collaboration were negative. "There was neither general nor specific, neither directed nor imitative correspondence. The strong probability is that no collaboration existed . . . While these content analyses did not *prove* the existence of noncollaboration (a feat logically impossible on the basis of available data), they did make it clear that a close, operating relationship between the German and Italian propaganda ministries was extremely unlikely"

(Berelson and de Grazia, p. 253). As a matter of fact, information secured later by occupation officials established the accuracy of this inference.

Different images of Stalin held by top members of the Soviet elite, as expressed in birthday speeches of December, 1949, may be interpreted to reveal something of their attitude toward him (Leites and others, 1951). Actually there was a sharp difference between the conceptions of the old Bolsheviks and those of other members.

	STALIN: LENIN'S PUPIL OR EQUAL?	
SPEAKERS	*Bolshevik Image*	*Popular Image*
Molotov, Malenkov, Beria	22	3
Shvernik, Voroshilov, Mikoyan, Andreyev	7	17
Bulganin, Kosygin, Khrushchev, Kaganovich	1	19

Although the communicators are much "closer" to the production of content than the "cultural patterns" of another section, nevertheless it is quite possible for inferences about them to be incorrect when based on the results of content analysis alone. Interpretations of motive are always difficult, and there is nothing in the procedure of content analysis which guarantees against mistaken interpretations. A single piece of content may be — and ordinarily is — consistent with more than one hypothesis about the characteristics or intentions of the communicator. There have been few studies which have attempted empirically to establish the relationship between the communicator and the content on the basis of independent material, i.e., to test the inferences based upon content analysis.

To determine the psychological state of persons and groups. A basis for inferences dealing with the psychological state of individuals and groups is provided by analysis of symbol materials produced *by* the persons about whom inferences are drawn. (This should be clearly distinguished from the analysis of symbol materials designed *for* the persons about whom inferences are drawn.) This involves the analysis of the social work or the therapeutic interview, of personality tests, of fictional creations, or of letters, diaries, or other personal documents.

In recent years there has been an increase in the application of quantitative methods to the analysis of psychological documents. According to a review of the literature in this field, a study quantifying "the structure of the personality of *one individual* life on the basis of a collection of personal documents . . . is of considerable importance . . . It sets forth a new conception of a 'population' for statistics: a population of events and traits within the boundaries of one person. If the initial promise of this method is fulfilled it will supply an important bridge between the statistical and the clinical points of view" (Allport, 1942, pp. 35-36).

One clear application — although the quantification is not always systematic, or even explicitly present — is in the field of projective tests (Rorschach, TAT, word association, etc.). The personality traits of the subject are inferred from the communications produced by them in response to ambiguous stimuli (an ink blot, a picture, an incomplete sentence). The attention given by the subject to form rather than color, or to the whole rather than the separate parts, or to aggressive rather than peaceful plots provides the analyst with the data on the basis of which he characterizes the subject in psychological terms. Usually only the presence or the absence of the content score is necessary to support the inference, but on occasion quantities are also brought into play. Thus it can be significant if the subject returns to certain topics (e.g., sexual references), or if he sees the same things in different stimuli, or if he elaborates at some length on a particular matter (for example, see Lindner, 1950). In any case, a psychological description of the subject is based upon analysis of the communications produced by him. The reader may obtain an excellent view of systems of content analysis currently applied to the Thematic Apperception Test in a recent volume edited by Shneidman (1951). Some impression of the use of content analysis with other projective techniques may be secured by consulting Abt and Bellak (1950) or Anderson and Anderson (1951).

An early suggestion for objectifying the psychoanalytic interview through content analysis listed a representative set of categories into which the content might be classified and "which it is predicted will aid in the orderly comparison of interview data" (Lasswell, 1938, p. 199). The broad headings of this proposal are:

$$\text{Pro- or anti-}\begin{cases}\text{Self}\\\text{Interviewer}\\\text{Other}\end{cases}\quad\text{by}\begin{cases}\text{Self}\\\text{Interviewer}\\\text{Other}\end{cases}$$

Some of the possible combinations are illustrated by these formulations:

"I am proud of my record" — Pro-self by self

"You insulted me" — Anti-self by interviewer

"He said you were a fraud" — Anti-interviewer by other

"He loved her" — Pro-other by other

For psychiatric purposes, the basic classification is thus stated in terms of self and nonself, approval and disapproval. In addition, refinements are suggested by forms of indulgence and deprivation and by patterns of internalized and externalized action. "Interviewers grow accustomed to expect characteristic constellations when there are strongly marked peculiarities in the associative style of the subject" (Lasswell, p. 204). No application of these particular categories to psychoanalytic interviews has been published, and thus their utility remains an open question. At the same time, it is clear that many psychoanalytic propositions are based upon data which can be analyzed by content analysis.

Less formal (and usually nonquantitative) studies of fictional materials to reconstruct the personality structure of the author have been based primarily on psychoanalytic categories. Assuming that the writer cannot fail to express deepest feelings in his art creations, the analyst proceeds to uncover them by attention to allegedly revealing particulars in the fictional works themselves.

The novels of Charlotte Brontë show evidence of an Oedipus theme; the chief concern of her sister Emily was with the emotional relations of children. Charlotte's novels, with their decrease of subjectivity and reduction of conflict, indicate a process of emotional integration which Emily did not achieve (McCurdy, 1947). The latent fascism of Knut Hamsun was betrayed by the authoritarianism of his novels, the acceptance of natural forces as unchangeable, the combination of brutality and sentimentalism, the motherly and housewifely role ascribed to women, the importance of the law of rhythm (Lowenthal, 1949). The novels of D. H. Lawrence reveal three stages of sexual development — Oedipal relations to his mother in the early novels, then homosexual relations, and finally the achievement of heterosexual relations (McCurdy, 1939). Dostoevsky's treatment of women in his novels reveals his oral character structure and his aggressiveness toward women (Kanzer, 1948). Charles Kingsley's *Water Babies* constitutes a renunciation of the physiological process of procreation and indicates his need for purification (Deutsch, 1947).

Occasionally such analysis has been applied to single poems (as in the Wormhoudt, 1949, investigation of the unconscious sources used by the romantic poets in England) and even to the explanation of the motives of artistic characters (e.g., Jones, 1947, on Hamlet; and Kris, 1948, on Prince Hal). Of such investigations, students of literature point out the invalidity of treating the work of art in any but its own terms — i.e., in terms of its own requirements and forms, rather than the author's biography — as well as the tendency to overlook the effect of the social milieu and literary traditions of the times. Content analysts point out the unsystematic and highly selective nature of the evidence usually presented to support the central hypotheses, and the absence of a sensible comparative control. Occasionally, however, such studies do develop stimulating insights which may enrich one's appreciation of the total art product of an author.

Some analytic applications have been made directly to psychological materials. For example, one study in "personal structure analysis" classified a series of letters written by a single person into several topical (e.g., self, money, death, religion) and attitudinal (e.g., moral, favorable, lazy) categories, and investigated their frequencies and particularly their contiguities (Baldwin, 1942). The resulting data indicated certain correspondences which allegedly represented the principal ideational clusters in the subject's life. In validating this quantitative analysis, by comparing its results with the impressionistic interpretations of trained psychologists, the author concludes: "If used in conjunction with other information about the case, the method confirms judgments already formed, adds some new insights, and presents problems whose solution leads to further insight. Even if used by itself, the table of correspondences is able to communicate considerable information to an observer" (p. 181). Thus this method of analysis not only "offers a technique for studying the unique pattern of the individual personality, an aspect heretofore not subject to objective analysis, (but) it points the way to a quantitative, systematic, objective method of clinical interpretation" (p. 183). The same basic method, though with certain statistical refinements, was applied to the analysis of the personal diary of a girl adolescent (Andrews and Muhlhan, 1943). From a frequency-contingency table including some twenty categories, the authors matched those frequency patterns with high degrees of correspondence. These groups were then taken to represent "congruent idea patterns" — a goal-identification factor, an interest-world factor, a socio-sexual immaturity factor, and so on. "We do not believe the technique to be of such power that other aspects of analyzed material should be ignored, but we do believe that the technique offers a device through which new and important insights into the hidden meaning of personality structures can be obtained" (p. 109).

This procedure is not limited to the analysis of personal documents in this exclusive sense; it has been applied, for example, to autobiographic materials of some literary standing. Thus Richard Wright's book *Black Boy,* an eloquent statement of the growth of a sensitive Negro boy in the American South, has been subjected to a detailed value analysis which attempts to systematize and categorize the picture of the world presented (White, 1947). Presumably under the tacit pressure of justifying the application of such methods to this material, the author actually organizes his report in terms of the findings which support the impressions of reading and those which modify them. Among the first, for instance, are the high frustration-satisfaction ratio and Wright's hostility to his father and ambivalence toward his mother. Among the latter are the primary emphasis upon physical safety, the lack of emphasis upon social goals, and the large amount of 100% disapproval of the people around him. In most studies of this kind, dealing with a relatively limited body of material, analyses of such a precise, quantitative, categorical nature probably add little to attentive reading and nothing to careful study. In most cases, the formal content analysis is not worth it.

Finally, the method has been useful in substituting systematic and objective data for "impressions" and "interpretations." One example is the application to the analysis of group interactions, mentioned above (Bales, 1950). Another is the application to the case records of social work clients.

The change in tension level throughout the period of a client's treatment was charted in one case on the basis of the "DRQ," that is, the discomfort-relief quotient (Dollard and Mowrer, 1947). This quotient is simply the figure resulting from the division of the number of discomfort words by the number of discomfort plus relief words combined, computed at different stages of the case. With a technique of this general kind, it may be possible not only to identify the specific points of high or low tension on the part of the client but also to measure the general trend in his tension levels from the beginning to the end of the case. In a similar case, social work records were analyzed primarily in order to compare the problems presented at different offices of the same agency (Berelson, 1947). The case records were classified in some 80 distinct sets of categories referring to various aspects of the "causes," conditions, and consequences of the case. From this analysis, for example, it appeared that the more intensely the client feels about his problem, the more likely he is to despair of a solution. In the same way, the analysis established a positive correlation between the client's attitude toward agency service and his degree of confidence in the solution of his problem.

From experience with this use of content analysis, it appears that the procedure is more difficult and time-consuming when "deeper" or more complicated symbol materials and meanings are under analysis, and thus in many cases is not worth doing. That the social work case record has been usefully treated in this manner, most students would agree (and so would many social workers). But there would be less agreement about the value of studies of subtle and complex psychological problems through quantitative symbol analysis.

To detect the existence of propaganda (primarily for legal purposes). The analysis of propaganda *techniques,* wherever they occur, is one thing: the analysis of the *existence* of propaganda is another. This application of content analysis, particularly associated with Harold Lasswell, was used during World War II by the Department of Justice in its investigations of allegedly subversive organizations and their publications.

At bottom, the procedure here is quite simple. Comparison of the content under study with that of other content explicitly identified with a certain propaganda position, or in some cases with explicitly formulated standards, serves to reveal the existence of propaganda. During the war, for example, parallels between allegedly native-fascist propaganda and the output of the Nazis were admitted in evidence by the courts in connection with sedition trials, and they contributed to convictions. At the same time, "many 'suspect' papers and radio stations were absolved from suspicion by the results obtained when issues and broadcasts were quantitatively analyzed. They were protected from the all too frequent practice of condemning an entire publication on the basis of a few colorful examples of bias in news handling or editorial comment"

(Lasswell, 1946, p. 76). In the same vein, the admission of content analysis data in courts of law is said to stem from "the possibility of error arising out of . . . assumptions and subjective reactions and predispositions . . ." (*Content Analysis,* 1948, p. 911).

This application of content analysis is based upon several "tests" described in Lasswell, 1949 (pp. 177-178):

Test	*Explanation*	*Illustration*
Avowal	"Explicit identification with one side of a controversy"	Of 243 self-advertisements of the *Moscow News,* about half referred to the authoritativeness of the publication
Parallel	"The content of a given channel is compared with the content of a known propaganda channel" (by themes)	In Pelley's publication *The Galilean,* there were 1195 statements consistent with Nazi themes and only 45 against them
Consistency	"The consistency of a stream of communication with the declared propaganda aims of a party to the controversy"	The percentages of editorials and articles in the Auhagen publications *Today's Challenge* and *Forum Observer* consistent with the Nazi aim "to exposè the weakness and decadence of the enemies of Germany" were 71% and 57%, respectively (as compared with 6% for the *Reader's Digest* and 10% for the *Saturday Evening Post*)
Presentation	"The balance of favorable and unfavorable treatment given to each symbol (and statement) in controversy"	In the *Moscow News* there were 1903 items favorable to the U.S.S.R. and 3 unfavorable, 69 favorable to other countries and 328 unfavorable
Source	"Relatively heavy reliance upon one party to a controversy for material"	About 20% of the material in the *Moscow News* came from acknowledged governmental and party sources
Concealed source	"The use of one party to a controversy as a source, without disclosure"	(No general evidence)
Distinctiveness	"The use of vocabulary peculiar to one side of a controversy"	Among the most frequently used political terms in the *Moscow News* were such distinctive symbols as commissariat, Soviets, masses, Stakhanovism, etc.
Distortion	"Persistent modification of statements on a common topic in a direction favorable to one side of a controversy"	Of the items in *Transocean* but not in the *New York Times,* 58% were consistent with Nazi views and only 3% were contradictory. Of the *Transocean* omissions, only 16% were consistent and 48% were contradictory

Such tests, then, state explicitly the characteristics which are used to detect the presence of propaganda.

This application of content analysis is not without its critics and questioners. On the legal side, the problems of clarity, reliability, subjectivity, and significance are raised (*Content Analysis,* 1948). The logical difficulties inherent in distinguishing correspondence (in content) from collaboration (in intent) are formidable, and there may be serious misinterpretations in getting from one to the other (on this problem, see Berelson and de Grazia, 1947).

To secure political and military intelligence. Stimulated particularly by the requirements of World War II, there has been a good deal of effort devoted to securing intelligence data of various kinds from the study of communication content. How is the enemy population reacting to our propaganda? What is the state of their morale? What are the enemy's expectations as to our next military moves? What are his evaluations of our military situation? What will he do next? Answers to such questions, often on quite crucial political and military matters, have been sought by means of content analysis.

Strong claims have been made for this function of content analysis: "By analyzing the content of what appears in the channels of communication in totalitarian powers, we may improve our predictions about the next move to be made by the controllers of policy in those states. Already some spectacular successes have been claimed by students of communication. Some of the analysts connected with the British Broadcasting Corporation have reported that they were successful in making certain predictions about the military plans of Germany" (Lasswell, 1942a, pp. 110-111). However, until very recently — until the forthcoming impressive review of this function of content analysis by Alexander George, which deals with the wartime experience of the Analysis Section of the Foreign Broadcast Intelligence Service of the Federal Communications Commission — there has been little scientific scrutiny of this whole field.

The record of governmental work during World War II, as analyzed by George, contains a number of successful predictions of major military actions. Perhaps the most dramatic of all was the case of the British predictions about the German V-weapon. At a time (fall, 1943) when there was some doubt whether the Germans had a secret weapon and, if so, what it was, the British analysts made a series of remarkably accurate inferences not only specifying the possession of the weapon but also its general type and even a succession of dates crucial in its development and use. The bases for their analysis were the generalizations provided by prior studies that ". . . German (home) propaganda never deliberately misled the German people in questions involving an increase of German power" and the reasoned assumption that ". . . Goebbels would be careful not to give the German public a promise of retaliation too far ahead of the date on which the promise could be fulfilled" (lest it boomerang into popular disillusionment). On the other occasions, content analysts correctly inferred that the Germans did not plan to move into Spain and Spanish Morocco in March, 1943; that they did not plan a major offensive on the Russian front in 1943; that they planned to withdraw from the Kuban bridgehead on the Eastern front; that they did not expect a high level of success in submarine warfare in 1943.

AUDIENCE OF CONTENT

To reflect attitudes, interests, and values ("cultural patterns") of population groups. Content analysis has been used to get at "the spirit of the age." On the assumption that communication content "expresses" or "reflects" some sort of *Zeitgeist,* some students have used it as an index to attitudes, interests, values, mores of the population. Thus the problem of discovering the topics people were interested in during a certain period or what they thought about various matters is studied by determining the nature of the content of popular communications at the time, and changes in such cultural patterns are studied in the same way.

There are two characteristics of this use of content analysis which need to be clearly recognized. The first is that inferences about population groups are made on the basis of content produced *for* them and not *by* them. Thus cultural patterns are derived from analysis of communications widely distributed in the community. The second point is that such problems are investigated in this way — i.e., indirectly — precisely because it is not possible to attack them directly. If we wish to study attitudes and interests and values today, we have certain scientific tools reasonably adequate for the job — or at any rate, more adequate than this second-level approach. But we cannot use them to study a historical people or a foreign people from whom we are barred; for such groups the indirect approach is necessary. However, the fact that it is the only way to get at the problem should not obscure the other fact that it is at best a substitute for the real thing. Both these points, incidentally, raise problems of validity to which we shall return later. The reader will find in Chapter 26 a discussion of the difficulties encountered in the study of national character, many of which are related to the present topic.

There are any number of examples of such studies in the literature. The "changing social attitudes and interests" in the United States from 1900-1930 were identified through selected analyses of popular magazines (Hart, 1933). For example, the general decline in the status of religion and religious sanctions during this period is documented not only by the amount of attention given to the topic but also by the decline of favorable references to organized religion. Such trends in the treatment of religion in popular magazines are taken to reflect a general weakening in the position of religion among the population at large.

A shift in popular concern, during this century, from the "idols of production" to the "idols of consumption" is allegedly indicated by the distribution of the subjects of biographical sketches in certain popular magazines (Lowenthal, 1943). Not only was there an increase in entertainers at the expense of politicians, but the entertainers themselves were drawn less and less from the serious arts. Substantially the same procedure has been used to infer the state of public morale from the communications designed by propagandists who have both totalitarian control of the media and, presumably, accurate and direct knowledge about morale itself. "Enemy communications are continually used for the purpose of estimating the state of morale among enemy peoples. This is the study of content as a basis of inference about the changing state of attitude toward war, and especially the intensity of the determination to carry on to victory" (Lasswell, 1946, p. 80). However, despite the great amount of work done on wartime propaganda, no quantitative index based upon formal categories has been devised which will reliably reflect popular morale when applied to propaganda materials. Totalitarian propaganda has also been used as a basis for the construction of a certain personality syndrome (the "compulsive character" of psychoanalytic theory) as a "distinctive type of character structure in the Nazi variant of German culture" (Kecskemeti and Leites, 1947, 1948a, b). Such psychological "traits" as revolt and submission, order and disorder, guilt and self-righteousness, doubt and dogmatism were documented from the propaganda material on the assumption that it could validly be used as an index of character structure.

Such studies, then, employ communication content as an index to various cultural patterns which they find "reflected" in the content. But how valid is this procedure? Can communication content properly be taken as a popular mirror of the time? Certain cautions on this application of content analysis must be mentioned.

In the first place, it is important to be completely clear about just *what* is being "reflected" in the communication content. Is it the generalized cultural and psychological level of the total adult population which somehow finds expression through the media? Or do the inferences from content refer only to the audience to the particular communications under analysis? If the latter, then the problem of changes in the composition of the audience in trend studies of content must be faced: that is, the changes in content may reflect changes in audience membership rather than changes in psychological properties of constant population groups. Along the same line, do the results actually refer to the psychological categories proposed or to some other less subtle aspects of reality? For example, does the change in the biographical subjects of popular magazines reflect a change in the public's attitude toward popular heroes any more than it does a change in the occupational distribution during a period in which commercialized sports and entertainment became big business in this country (in which case content data are much less satisfactory than census data on occupations)? Such questions must be answered in the particular study for clarity not only of interpretation but also of the analysis itself.

The basic rationale for the use of content data to permit noncontent inferences of this kind has been mentioned in the literature. Here, for example, are two such arguments:

> The common day-dreams of a culture are in part the sources, in part the products of its popular myths, stories, plays, and films. Where these productions gain the sympathetic response of a wide audience, it is likely that their producers have tapped within themselves the reservoir of common day-dreams. Corresponding day-dreams, imperfectly formed and only partially conscious, are evoked in the audience and given more definite shape. (Wolfenstein and Leites, 1950, p. 13)

> Our first assumption in this study is that popular drama can be regarded as a case of "social fantasy" — that the psychological constellations in the dramatic work indicate sensitive areas in the personalities of those for whom the work has appeal; their needs, assumptions, and values are expressed ("projected") in the drama. The successful play must be attuned to the audience. There can be no claims that dramatic material reflects the total personality of the individual who enjoys it. (McGranahan and Wayne, 1948, p. 430)

Thus this use of content analysis assumes that the nature of the audience about which the inferences are made is at least *a*, and perhaps *the*, major determinant of the content which is being analyzed.

But this is clearly not always, or perhaps not usually, the case. Content is produced by particular agents (writers, producers, directors, editors, etc.) and it is often *their* conception of audience attitudes and values, biased as it may be, which determines what appears. Not only that; communication content is also subject to the pressure of special interest groups which may have enough weight to determine, or at least influence, the nature of the content.

For example, a behind-the-scenes campaign by the Catholic Church to "purify" popular publications on sex issues could have effected a major change in the appearance of such material with no corresponding and certainly no prior changes in popular attitudes. This kind of limitation is recognized by some writers in this area (e.g., Hart, 1933, p. 386) and directly argued by others:

> In part the recurrent patterns . . . may be consciously contrived . . . But both the positive sanctions of the box office and the negative ones of the code are probably indicative of more or less widely diffused feelings and attitudes in American culture. In part also, the selection of certain themes and omission of others would seem to be less calculated, expressive of deeper and less conscious emotional tendencies. Within the limit of a code or a general plot formula there remains the choice of a considerable range of possibilities. (Wolfenstein and Leites, 1950, p. 14)

For several reasons, then, the mechanical application of content analysis to problems of this kind should be avoided. It is difficult to know under what conditions inferences can be validly drawn about the total population or only about the particular audiences; to what extent they refer to the audience proper or to the producer's conception of the audience or to the producers themselves as (atypical) members of the audience; whether they legitimately apply to psychological variables in this sense; whether they are correctly based upon a conception of audience characteristics as a source of the content or as an effect of it or both; whether popular values are somehow identifiable, however roughly, as a direct quantitative function of content emphases. In short, the whole relationship between the content and audience characteristics allegedly "reflected" in it is far from clear. At the least it is far from a one-to-one relationship, and this fact in itself is a cautionary note against the over-interpretation of content data.

EFFECTS OF CONTENT

To reveal the focus of attention. Among the clearest effects of the mass media in modern society has been their extension of people's attention areas. Whereas in simpler societies people were limited to matters which came to their or their neighbors' direct attention, the proliferation of technical means of communication has resulted in bringing the whole world into the ordinary man's focus of attention. In a sense, every content study is an analysis of what comes to people's attention. Some studies, however, are more or less limited to this use of content analysis. They have dealt with different media, with widely differing problems, and with different analytic methods.

One simple form of such a study, particularly associated with Lasswell, is illustrated in his *World Attention Survey,* which deals with the occurrence of selected political *symbols* in various important newspapers. Analysis of the relative frequency of such symbols reveals, it is claimed, that "the stream of public attention is related to policy . . . In the *Voelkischer Beobachter* . . . we see less and less attention paid to the Soviet Union, previously a target of bitter hostility. This indicated that Germany was clearing the path for a sudden change in diplomatic orientation, as was learned when the (Soviet-German) pact was announced" (Lasswell, 1941, p. 457).

Another form of content analysis has to do with the *subject matter* distribution of what appeared in the selected media of communication. For example, the average percentage of foreign news in American morning newspapers in 1927 was shown to range from a low of about 2% to a high of about 9% (Woodward, 1930). Even the personal conversations of a group of people were distributed in subject-matter categories in order to identify the major differences in what men and women talk about: the former concentrated upon business affairs and the latter upon personal affairs (gossip) and clothes (Landis and Burt, 1924).

Such analysis is not limited simply to the relative distribution of different symbols or subject matters. Some studies have added *direction* (i.e., approval or disapproval) of the material as well. For example, the relative approval or disapproval of various educational topics in newspaper editorials indicates the areas in which public education is given most public praise and public criticism. Thus although the public's responsibility for supporting education is almost everywhere accepted, there was widespread disapproval of costs involved (Foster, 1938).

	PERCENTAGE OF EDITORIAL COMMENT		
	Favorable	*Neutral*	*Unfavorable*
Public responsibility for supporting education	90	8	2
Cost of education	8	15	77

Another relevant category in the study of attention areas is the *source* of communication material. For example, a study of the origin of war news in the *New York Times* during an early period of the first World War revealed that the Central Powers supplied less than five percent of the front page stories as against fully seventy percent supplied by the Entente Powers (Foster, 1937). Thus there was nothing like a balance in the war news provided American readers in terms of the sources of the material coming to their attention.

Content analysis has also been applied to the determination of attention areas in fictional materials in the mass media. For example, the major characters in popular motion pictures have been described by these procedures. About 46% of the major characters in feature films are either wealthy or well-to-do, and only about 17% of them are poor or destitute (Jones, 1942). Similarly, about 68% of the characters wanted love as a major value as against only about 9% who wanted "rightness" or some other social value. And almost two-thirds of the characters have all of their desires fulfilled at the end of the picture. A more intensive and a more imaginative analysis of motion pictures included such categories as the general themes of the films, personal characteristics of the cast, kinds of clothes, circumstances of meeting and lovemaking, crime, vulgarity, recreations, and goals (Dale, 1935). In the same way, the world as it appears in the daytime serial has been described in great detail. Here again, for example, the major characters enjoy a high occupational status (much higher than that of the typical listener) and the major themes deal overwhelmingly with problems of personal relations rather than public affairs (Arnheim, 1943).

In any general analysis of the role of communications in forming people's opinions and attitudes, it is important to stress the factor of attention. The self-selection of what will be read or listened to means that attention is not automatic and hence not always in correspondence with communication content. In this kind of investigation, content analysis plays its part in presenting an objective and systematic picture of the communication content which is *available* for public attention, whatever the actual exposure.

To describe attitudinal and behavioral responses to communications. The culminating factor in the study of communications is their effect upon individuals, groups, institutions, society. Effects can be short-run or long-run; they can be immediate or remote; they can reinforce or convert; they can deal with information, knowledge, opinions, taste, skill, "behavior." Under certain conditions, content analysis can contribute to the investigation of effects.

There are three ways in which content analysis has been employed in the study of effects. The first involves the analysis of communications which are themselves expressions of the effect of previous communications. In other words, the analysis deals directly with the material as response, instead of, as in most cases, with communications as stimuli.

Although there are relatively few examples of this type of effect analysis, it is worth noting as a possibility for further exploration. Two cases in which it was applied dealt with the reception of scientific work by the intellectual community. The acceptance of Freud's dream theory, as an instance of his total contribution, was a good deal less than general in the 350-or-so psychology textbooks analyzed (Herma, Kris, and Shor, 1943). And the writers who accepted the theory identified it as Freud's less than those who rejected it. In the other case, the reception of *The American Soldier* in scholarly and popular journals was analyzed into a typology of "affective responses" descriptive of the scholarly images and allegiances elicited by that work (Lerner, 1950). The reason there are not more such direct content analyses of effect is obvious: except in such limited forms as textbooks and book reviews, the response material is simply not available.

The second type has also had limited application as yet, largely because of technical difficulties involved in its use. It is the direct correlation of content data and response data, i.e., studies in which the frequencies with which certain categories appear in the content are related statistically to the frequencies of response in the audience. Such a relationship was attempted in an investigation of the effect of the content emphasis of selected political arguments upon the recognition and the acceptance of the arguments (Berelson, 1942). Although frequency in the content did determine to a large extent the recognition of the arguments, it was by no means the major influence upon agreement. The correlation between content emphasis on the arguments and agreement with them was small because of the role of predispositions in determining agreement. However, it was demonstrated that content emphasis decreases indecision and that it tends to regulate relative agreement among arguments "within" the influence of predispositions.

Thirdly, there is the direct inference of effects from content data, without reference to data on response itself. This is probably the most frequent application of content analysis to effect problems. In fact, there are a large number of content analyses which simply rephrase their results in effect terms — mistakenly, in view of the role of predispositions and other factors involved in the situation. More sophisticated studies of content interpret responses against the psychological characteristics of the audience. For example, the effectiveness of Communist propaganda in Chicago during the depression of the early 1930's was evaluated on the basis of a systematic analysis of its themes and appeals (Lasswell and Blumenstock, 1939). The authors point out the relatively high proportion of key symbols which were distinctive to Communist propaganda; "thus the nucleus of Communist propaganda was composed of symbols which were counter-mores in America" (p. 352). As time went on there was a declining emphasis upon certain distinctive symbols and thus a somewhat greater degree of adaptiveness in Communist propaganda which presumably increased its effectiveness. Again the content of Communist propaganda was not particularly skillful in combating nationalistic and individualist symbols. "Despite some favorable circumstances, Communist propaganda was blocked by American nationalism and individualism and thus contributed more to the catharsis of insecurities than to the readjustment of American symbols and practices. It made only a modest contribution even to the appearance in America of the revolutionary pattern of the epoch" (p. 357).

Here the content analysis of Communist propaganda provided the basis for a psychological interpretation of its effect upon the recipients. This same method was used in a "sociological" analysis of the effectiveness of the round-the-clock bond drives by Kate Smith during World War II (Merton, 1946) which puts the content data in a framework of response data. Fully half of the Smith propaganda to promote sales was composed of the sacrifice theme (sacrifice on the part of the serviceman, the civilian, herself); and the interpretation of Smith's effectiveness derives partly from the prominence of that theme. "The triangulation of sacrifice . . . developed in many listeners a strong sense of unworthiness and guilt. They felt compelled to do something more to keep self-esteem . . . After the purchase, there was strong evidence of catharsis. Duty had been done and guilt swept away" (pp. 54-55).

In a psychological context, finally, the effect upon third-grade school children is implicit in an analysis of the values presented in their readers (Child, Potter, and Levine, 1946). The types of behavior rewarded in the books is assumed to have a guiding effect upon the behavior of the child, who may model himself, consciously or unconsciously, upon this material.

Mode of acquisition	*Percentage in which rewarded*
Effort	93
Buying, selling, trading	80
Asking, wishing, taking what is offered	68
Dominance, aggression, stealing, trickery	41

Just as the effect area in communication research is of major importance (see Chapter 28), so this application takes on central importance in the development of content analysis. More refined procedures and careful establishment of relationship between content and the specific type of effect under investigation are necessary in order to achieve desirable advances in the scientific study of the effect of communications.

THE UNITS OF CONTENT ANALYSIS

By definition, content analysis calls for the quantification of content elements. As a basis for counting relevant aspects of the content, certain standard subdivisions of it must be used.

Before specific units are discussed, three technical distinctions must be mentioned. *First* is the distinction between the recording unit and the context unit. The recording unit is "the smallest body of content in which the appearance of a reference is counted" (a reference is a single occurrence of a content element). The context unit is "the largest body of content that may be examined in characterizing a recording unit." For example, the recording unit may be a single word or term, but in order to note whether the particular term was treated favorably or unfavorably the analyst will be instructed to consider the entire sentence in which the term appears (the context unit). Or a sentence may be the recording unit and the paragraph the context unit; thus the paragraph would be taken into account in coding the sentence (Lasswell, 1942, p. 16). *Second* is the distinction between the unit of classification and the unit of enumeration. This involves simply the distinction between the basis on which the content is analyzed or classified and the basis on which it is tabulated or reported. Thus an analyst may classify an entire news story into a given category, and then report the result in terms of column inches. Here the news item is the unit of classification and the column inch is the unit of enumeration (Woodward, 1934). *Third* is the related distinction between levels of analysis, in which a prior analysis of a subsidiary unit characterizes the larger (parent) unit for later analysis. For example, classification of sentences may serve to characterize political articles as "biased," defined as meaning greater than 75-25 distribution of support to one side. Once the sentence analysis was completed, the articles themselves could be quantified for comparisons by source, channel, etc. Here the preliminary analysis by sentence provides the basis for the subsequent classification by article.

In the literature of content analysis five major units of analysis have appeared (for verbal symbols): words, themes, characters, items, and space-and-time measures.

WORD

The smallest unit generally applied in content analysis is the word. This unit (which includes word compounds like phrases as well as single words) is identical with the "symbol" or "unit symbol" of Lasswell (1942) and the "term" of Leites and Pool (1942). Its application results in a list of relative frequencies of selected words or selected categories of words.

This unit has been employed most extensively in three sets of studies. The first of these is the political symbol analysis associated with Lasswell. In his *World Attention Survey,* for example, he studied the relative occurrence in the world press of a list of "key symbols" in modern politics: freedom, liberty, constitutional government, fascism, national socialism, communism, and so on (Lasswell, 1941, 1942). The second application of the word (or word-compound) as a unit is made in some analyses of literary style. Thus, a series of studies report the major terms employed by a wide range of English poets (Miles, 1951). The third important use of word units in content analysis occurs in the studies of readability, to determine the ease or difficulty with which communication material can be read and comprehended. Such studies construct readability formulas on the basis of personal references, prepositional phrases, different hard words, words with prefixes and suffixes, and so on (Gray and Leary, 1935; Lorge, 1944; Dale and Chall, 1948; Flesch, 1951).

THEME

The next larger unit of content analysis is the theme. In its most compact form, the theme is a simple sentence, i.e., subject and predicate. In other words, a theme is an assertion about a subject matter. Thus, it is a sentence (or sentence-compound), usually a summary or abstracted sentence, under which a wide range of specific formulations can be subsumed. Again, the theme has been given various names in the analytic literature: "assertion" and "statement" by Leites and Pool (1942); "proposition" by Lasswell (1942); "idea" by Kris and Speier (1944); "issue" and "argument" by Waples and Berelson (1941); and "thema" by Child, Potter, and Levine (1946). The specific nature of analytic themes will depend, of course, upon the particular content and problem at hand.

Thematic analysis has been done on both political and nonpolitical materials. Analyses of Axis propaganda in World War II identified such themes as "The U.S. is internally corrupt" (Lasswell, 1949) or, more specifically, "England does not fight herself, but gets her Allies to do her fighting" (Berelson and de Grazia, 1947). A set of themes describing "social optimism" (i.e., emphasis on social rather than individual responsibility) and "social pessimism" (i.e., affirmation of man's need for mysticism) was analyzed in Protestant sermons (Hamilton, 1942). The content of Kate Smith's broadcasts on a marathon bond drive was also analyzed by themes — the themes of sacrifice, participation, competition, facilitation, and familial and personal appeals (Merton, 1946).

The theme is among the most useful units of content analysis, particularly for the study of the effect of communications upon public opinion, because it takes the form in which issues and attitudes are usually discussed. But it is at the same time among the most difficult units of analysis from the standpoint of reliability, especially if it is at all complicated (i.e., more than just a simple sentence). One elaborate, complex method of attacking this problem is to break the theme down into its components, to analyze by them, and then to reconstruct the theme by a "mechanical" process (Waples and Berelson, 1941). Thus the theme, "Roosevelt's New Deal, by having the right man in the right position, has saved the farmer and the working man from ruin," was analyzed into its several (relevant) components:

SUBJECT MATTER:	Record of the Democratic administration, domestic policy, general
Objects:	labor (workers), farmers
Values:	economic security, high indulgence
Method:	personnel selection

These components, rather than the theme itself, were coded on punch-cards and the themes were later reconstructed by running the appropriate combinations on a punch-card machine. It was believed that the detailed method of coding would be more reliable in this case. Such analysis has been called "atomic" as against the "molecular" analysis of the theme as a whole.

CHARACTER

Use of the fictional or the historical character as a unit is appropriate in analyses of stories, dramas, and biographical sketches. In cases where a narrative is woven around particular persons, they are the most convenient units on which to base an analysis. The entire story or sketch provides the basis for specific classifications, and the entire item must be read before any coding decisions are made. There are examples in the analysis of magazine fiction (Berelson and Salter, 1946), motion pictures (Jones, 1942), radio serials (Arnheim, 1943), and biographical articles (Lowenthal, 1943).

ITEM

The unit most frequently used in content analysis is the item, i.e., the whole "natural" unit employed by the producers of symbol material. The item differs among the different media: it may be a book, a magazine article or story, a speech, a radio program, a letter, a news story, an editorial, or any other self-contained expression. This unit classifies material in the large, so to speak, and for this reason it has been frequently employed, especially in the case of subject matter analyses for which it is particularly suitable. Analysis by the entire item is appropriate whenever the variations *within* the item are small or unimportant. Thus, for example, news stories can often be reliably classified on a subject matter basis (e.g., domestic politics, international affairs, crime, labor, etc.). But if detailed categories are added which introduce variations within items, the item unit may become inappropriate.

SPACE-AND-TIME MEASURES

Rather than analyze broadly by items, some studies have classified content by such physical divisions as the column inch or the page or the line or the paragraph (for print), or the minute (for radio), or the foot (for film), thus allowing for somewhat more precise distinctions than are possible with the item unit. Thus, a news story containing two subject matters may be assigned *in toto* to the dominant one under item analysis, whereas under this system, say, eight inches will be assigned to one topic and two inches to the other.

THE INTERRELATION OF UNITS

There is no reason, of course, why a particular study must use only one of the possible units of content analysis. The choice of the appropriate unit depends upon the problem and the content under investigation, and this may necessitate the use of different units within the same study. For example, in an analysis of magazine stories, both the character and the item were used to test different hypotheses about the content (Berelson and Salter, 1946). Again, in an analysis of letters the theme was used in one part of the study and the letter as a whole in another (Wyant and Herzog, 1941a, b).

What unit should be selected for a particular study? No general answer can be given. At this stage in the development of content analysis, the best suggestions seem to be these:

(a) More than one unit of analysis can profitably be employed in a single study.

(b) Word counts and large units (items and space-time measures) are most appropriate for studies of straight subject-matter emphases.

(c) Large units ordinarily provide as valid an analysis of direction (i.e., approval-disapproval) as small units.

(d) Large units take less coding time than small units, for the same categories and materials.

(e) Since themes add another dimension to subject matter, they are usually harder to analyze than symbols or large units, i.e., they take more time for the same reliability.

(f) The theme is appropriate for studies of meanings and of relationships among meanings.

The Categories of Content Analysis

Content analysis stands or falls by its categories. Particular studies have been productive to the extent that the categories were clearly formulated and well adapted to the problem and to the content. Content analysis studies done on a hit-or-miss basis, without clearly formulated problems for investigation and with vaguely drawn or poorly articulated categories, are almost certain to be of *indifferent* or low quality, as research productions. With the important caution that categories are most appropriately designed in terms of the particular problem under investigation, it should be useful to review the major types of categories which are found in the literature and illustrate their application to different kinds of material.

In this review, a rough distinction can be made between the *what* and the *how*, i.e., between substance and form. Some kinds of categories (e.g., subject matter) deal clearly with *what* is said and others (e.g., form of statement) with *how* it is said. Although the borderline is not always precise and distinct, it may be useful to keep this dichotomy in mind in considering the categories used in content analysis.

"WHAT IS SAID" CATEGORIES

Subject matter. This is perhaps the most general category used in content analysis studies and it answers the most elementary question: What is the communication about? This is the basic question in analyses primarily concerned with determining the relative emphases given to different topics in a body of communication content. Topics are subject-matter categories in the same sense that a subject of an active sentence is its subject matter. Subject matters are usually quite obvious in a body of communication content and hence easily analyzed. Subject-matter categories have been especially used in trend studies in various media, with the same categories analyzed at different points in time.

Direction. This category, also called "orientation" (Allport and Faden, 1940) and "character" (Harris and Lewis, 1948), refers to the *pro* or *con* treatment of a subject matter. Basically, the question is: Is the communication for or against the particular subject, or neutral toward it? Subdivisions of this general formulation are possible and some have been made, but the substance of the category is this "plus-ness" or "minus-ness" of the content toward a topic. The analysis by subject matter reports that 2% of the content deals with a given subject matter, and the analysis by direction reports how much of the 2% was pro, con, and neutral.

Standard. This category, also called "grounds," refers to the basis on which the classification by direction is made. It "characterizes the nature of indulgence-deprivation" (Lasswell, 1942) by asking in what terms the evaluation is made. The emphasis upon standards in this sense has mainly come from the Lasswell group, who have done most to develop the category.

The most general standards are two: strength-weakness and morality-immorality. Is the subject matter approved because it is strong or because it is moral? Within this dichotomy, however, several representative subdivisions have been made, including the categories safety, efficiency, power, goods, and respect under the heading of strength-weakness, and the categories truth, mercy, heroism, and loyalty under morality-immorality. In addition, other standards have been specified: propriety, divinity, legality, beauty, consistency, probability, euphoria, and omnibus (Lasswell, 1941; Kaplan and Goldsen, 1949).

Values. Closely related to standards is the category of values, also called "goals" (Berelson and Salter, 1946) and "wants" (Jones, 1942): what people want or get. This category is usually, though not always, found in analyses of fictional materials, applied to the characters. What are they after? People can be after various things — money, love, social position, career advancement, health — and these are their values.

Various subdivisions of the category have been made for different purposes. Representative values in Lasswellian analysis are income (money and/or material goods), safety (health or bodily integrity, life), deference (power, fame, reputation or prestige, rightness, love and affection). One entire system of content analysis is based upon values (White, 1951). Both psychological and political materials have been subjected to detailed analysis with the scheme, in its latest revision, of 29 specific "goals" under such headings as physiological, social, egoistic, fearful, playful, practical, and cognitive, and an additional 21 "standards of judgment." It is probably fair to say that rigid analysis by such systems of categories, unadapted to the particular problems at hand, may be uneconomic and relatively unyielding. The value category has been most useful when it has been formulated most specifically, i.e., "closer" to the actual content.

Methods. Values deal with the ends of action; methods deal with the means employed to realize the ends. How are objectives to be secured? This category has been used on political materials, in terms of such methods as analysis (diagnosis, prognosis, research), propaganda (content, quality, channels), negotiation, organization (personnel selection, administration, governmental regulations), use of economic goods" (bargaining, inducement), violence. More useful is the system of interaction categories devised for the analysis of group functioning, which can perhaps be considered as methods for reaching group decisions (Bales, 1950). See Chapter 10 for a detailed discussion of these categories.

Traits. This category, also called "abilities, subjective states" (Leites and Pool, 1942), includes ordinary personal characteristics, certain psychological traits, and other terms used in describing people. Although it is usually applied to persons, the category has also been used to characterize institutions and policies as well.

Many analyses of fictional materials have employed this trait category in one form or another. The composition of the fictional population of short stories, radio serials, motion pictures, and of biographical sketches in popular magazines has been described in such terms as age, sex, occupation, marital status, social class, religion, residence, and nationality (Jones, 1942; Arnheim, 1943; Lowenthal, 1943; Berelson and Salter, 1946; McGranahan and Wayne, 1948).

Actor. This category refers to the person, or group, or other object, which appears in a central position as the initiator of an action. Who is represented as undertaking certain acts?

Analysis of this category has demonstrated that the actual or alleged actors vary in response to other conditions. In one case a simple analysis of the relative frequencies with which "German forces," "Italian forces," or "Axis forces" were engaged in the actual fighting in the North African theater at a certain period in World War II produced a possible basis for the prediction of military action (Berelson and de Grazia, 1947).

Authority. This category, also called "source," refers to the person or group or object in whose name a statement is made. Some uses of the authority category have dealt simply with the specific sources of quotations; an unpublished analysis of the German radio during the early years of World War II, for example, disclosed the relative frequencies with which various American sources — persons, newspapers, radio stations, groups — were cited with approval. In addition, systematic analyses have been made of "impersonal" authorities; a study of wartime radio propaganda on a single event — the Altmark case — included such authority categories as universality, documentary evidence, neutral authority, and historical precedent (Bruner, 1941).

Origin. This category has been used in a few studies to identify the place of origin of the communication. Where did it come from? This is sometimes relevant in indicating how widely or narrowly the audience's attention is being directed. For example, a trend analysis of weekly country newspapers revealed a sharp shift toward localism in the point of origin of the content over a long period of time; the documentation was based upon a classification of origins — local, county, state, nation, foreign countries (Taeuber, 1932).

Target. This category, also called "addressee" (Waples and Berelson, 1941) refers to the group to whom the communication is particularly directed. It can be used to reveal the degree to which special groups rather than the general public are being addressed. Such phrases as "Labor will do well to remember that . . ." or "To American farmers I say . . ." illustrate the nature of the category. It has been used very seldom, and then without much return. In an intensive analysis of communications in a presidential campaign, its utility was so small that it was not even reported (Waples and Berelson, 1941).

"HOW IT IS SAID" CATEGORIES

Form or type of communication. This has to do with the ordinary distinctions among forms. Trend analyses of book production, for example, differentiate between fiction and nonfiction (Funke, 1937). Studies of radio content begin — and usually end — with the classification of material into classical music, popular music, drama, news, variety, talks, etc. (Albig, 1938; Bartlett, 1941; Baker, 1949). A few special studies have used language as a category; for example, an analysis of radio programs designed for foreign-language-speaking audiences classified content into English, German, Yiddish, Italian, and other languages (Arnheim and Bayne, 1941). Analysis of weekly newspapers distinguished between "news material" and "magazine material" (Willey, 1926). This category is obviously gross but it is often useful, particularly as a framework within which other distinctions can be made.

Form of statement. This category refers to the grammatical or syntactical form in which the communication is made or to its structural components. Because of this specificity of the category, it has usually been applied on a sentence-by-sentence basis.

In Lasswellian formulations, the standard classification for this category is threefold (Lasswell, 1942): fact statements (e.g., "The Communists want to rule the world."); preference statements (e.g., "The Communists ought to rule the world."); identification statements (e.g., "I am a Communist.").

Other forms of these basic categories can be distinguished. For example, fact statements can refer to the present or to the future (in which case they are expectation statements). Preference statements can be simple preferences or they can be formulated as imperatives for action (in which case they are demand statements). The generality of the categories and their lack of specific content have made them of relatively little use, except to reveal changes over a long period of time (Kris and Leites, 1947).

Intensity. This category, sometimes called "emotionalism" or "sentimentalization" (Waples and Berelson, 1941), refers to the "strength" or "excitement value" (Jacob, 1942) with which the communication is made. In many cases the analyst is interested in gauging the relative intensity of different contents, particularly because this factor is believed to be especially effective in influencing readers and listeners. Thus the "emotional component" of communication content is a relatively popular category for analysis. Despite interest in the category, attempts to devise generally satisfactory indicators have not been altogether successful. Although such methods have "worked" in specific cases (i.e., produced meaningful results), they have not been generalized into indicators of intensity applicable to a variety of problems in content analysis (Waples and Berelson, 1941; Jacob, 1942; Wright and Nelson, 1939).

Device. This category refers to the classification of a body of content on the basis of its rhetorical or "propagandistic" character. The category usually involves a whole complex of content, and is thus difficult to analyze objectively. Such propaganda "tricks" as name-calling, glittering generality, card-stacking, and bandwagon (Lee and Lee, 1939) are examples of this device category. Analysis of orations according to the type of argument used (e.g., argument by authority, argument from division, etc., as in Palmer, 1934) or according to the figures of speech employed (e.g., personification, rhetorical questions, metaphor, etc., as in Runion, 1936) also illustrate the category. In the field of literary analysis, this category is illustrated by the investigation of the frequency with which the pathetic fallacy was used by English poets over the period of a century (Miles, 1946).

TECHNICAL PROBLEMS

After the units and the categories of the analysts have been selected, there remain several technical problems to which attention must be given. Although they are all general problems of social science research, and not distinctive to this particular method, their application to content analysis presents some special problems.

WHEN TO COUNT CAREFULLY

Content analysis can be a time-consuming and expensive operation. Someone must go through the chore of carefully counting the occurrence of categories, on many of which difficult decisions must be made. This job is not always simple or easy, and although the real costs are usually obscured by the free labor, voluntary or enforced, of graduate students in the universities, they are nevertheless high. In research, as elsewhere, such costs must be reckoned against the returns.

Review of a number of so-called "qualitative" analyses suggests similarities and differences between them and so-called quantitative analyses (which are spelled out in detail in the author's book): (1) much "qualitative" analysis is quasi-quantitative; (2) "qualitative" analysis is often based upon the presence or absence of particular content rather than relative frequencies; (3) "qualitative" analysis is done on small or incomplete samples; (4) "qualitative" analysis usually contains a higher ratio of known content to content statements than quantitative analysis; (5) "qualitative" analysis is relatively less concerned with content as such than with the content as a "reflection" of "deeper" phenomena; (6) "qualitative" analysis employs less formalized categorization than quantitative analysis; (7) "qualitative" analysis utilizes more complex themes than quantitative analysis.

One may ask what are the conditions under which one should do a careful counting job. When can quantification add to a "qualitative" analysis? The following seem to be the conditions under which quantification is warranted.

1. *Count carefully when a high degree of precision and accuracy is required in the results.*

Here, as elsewhere, precise measurement is justified only when considerable precision and accuracy are needed in the results. In content analysis they may be needed to choose between alternatives, to "nail down" important conclusions, to settle impressionistic disagreements, to distinguish small differences, etc. Ordinarily this means that a comparison or other relationship is involved, implicitly or explicitly. But only as much precision should be bought as is necessary to the purposes of the study. It is the pretentious counting of the obvious or the unnecessary which gives content analysis a bad name.

2. *Count carefully when a high degree of objectivity is required in the results.* This is related to but not identical with the preceding maxim. Objectivity (i.e., the process of securing data which are independent of the character of the observer) is ordinarily prerequisite to the attainment of precision or accuracy. If the counting process is carried out with some care, the descriptive results can be accepted with a higher degree of confidence than those produced by impressionistic, subjective, nonsystematic procedures. It is true, of course, that under some conditions results may be spuriously objective. If a set of categories is so complex that it cannot be reliably analyzed, then the numerical values which purport to represent their incidence in the content are no better than impressionistic data — and perhaps they are somewhat worse, since they carry an aura of objectivity without the substance.

3. *Count carefully when the materials to be analyzed are representative enough to justify the effort.* If the materials under analysis are fortuitous, irregular, uneven, or otherwise unrepresentative of whatever universe of content they are meant to represent, then the effort of counting does not seem to be worth while. Any numerical results must be so partial and diluted that the additional margin of error introduced by impressionistic measurement is not critical. By and large, counting is not worth the costs in such circumstances, nor is it justified by the results.

4. *Count carefully when the materials to be analyzed are so many as to be unmanageable without such a summarizing procedure.* If the content is of such magnitude that it cannot be adequately handled without counting, then, of course, counting is necessary. For example, some trend studies of newspaper or radio content would be completely unmanageable without some form of counting, however crude. One mind simply could not review page after page of newspaper content and produce trend data of any consequence without the counting process.

5. *Count carefully when a high degree of specifications of categories is both possible and necessary or desirable.* The very business of counting demands that categories be precisely defined so that their incidence in the content can be correctly ascertained. If the study of the German film had been quantified (Kracauer, 1947), it would have required clear definitions of such content categories as "championing or even visualizing the cause of liberty," "testifying to the existence of a state of paralysis," "revealing the inner workings of the paralyzed collective soul," etc. If it is impossible or extremely difficult to specify categories sufficiently to allow careful counting of a reliable nature, then obviously, the results do not carry any guarantee of their validity.

6. *Count carefully when the categories appear with relatively large frequencies.* No matter how large the body of content, systematic counting may not be justified if the relevant categories occur only infrequently. In that case subjective procedures may be completely serviceable for the purpose at hand. For example, if one were concerned, as some Congressmen might be, with the degree to which American movies present left-wing political ideas, it would not be necessary to undertake a full-scale quantification, despite the large body of materials under analysis. This maxim is particularly applicable to those analyses which depend upon the presence or absence of the categories for their conclusions or interpretations.

7. *Count carefully when the content data are to be statistically related to numerical non-content data.* This is by no means always true, but it is often the case that counting is called for in order to provide numerical content data for correlation with other numerical data. This involves the use of content analysis for the construction of general propositions about communication and not simply for descriptive or historical purposes. For example, content trends on public issues might be correlated with attitude trends, or the intensity of the content with behavioral data, or the readability score with measurement of actual readership, or the "movement" in the content of social work cases with other indices of therapy. In none of these cases are numerical values for content *necessary* for establishing a relationship, but they may be more readily applied in such studies.

As in other scientific investigations, content analysis should be done not as precisely as possible but rather as imprecisely as possible —

that is, as roughly as the circumstances of the study will allow (in order to minimize costs relative to returns). This implies that under normal conditions careful counting should not be done unless it is quite necessary. It is probably not too strict to insist that all the above conditions be met, or rather all of the first six, since the seventh is a special case not always applicable. If the study does not deal with a *large* and *representative* body of material to be analyzed in terms of a set of highly *specifiable* categories which appear with *substantial frequencies,* in order to produce *objective* and *precise* results — if these conditions are not met, careful counting is probably not warranted.

THE PROBLEM OF RELIABILITY

By definition, content analysis must be objective. The analyst's subjectivity must be minimized in the effort to obtain an objective description of the communication content. The necessity to achieve objectivity, or near-objectivity, for content analysis raises the problem of reliability.

Reliability is used here in the standard statistical sense of measuring the degree to which the data are independent of the measurement instrument. Regardless of who does the analysis or when it is done, the same data should be secured under similar conditions. Thus the analysis of communication content rests upon two kinds of consistency: (1) consistency among analysts — that is, different coders should produce the same results when they apply the same set of categories to the same content; and (2) consistency through time — that is, a single coder or a group of coders should produce the same results when they apply the same set of categories to the same content but at different times. The acceptance of content analysis as an instrument for scientific research depends to a large extent upon the achievement of high reliability on both counts. What has been the record of content analysis in this regard? How reliable has it been in actual studies and in specially devised reliability experiments?

First, whatever the actual state of reliability in content analysis, the published record is less than satisfactory. Only about 15-20% of the studies report the reliability of the analysis contained in them. In addition, about five reliability experiments are reported in the literature. In brief, then, the simple fact is that the whole problem of reliability in content analysis has been given insufficient attention for the proper development of the field. The reliability of a research technique can hardly be established on the basis of reports from only one out of every five or six studies. Nor does the situation appear to be improving: there are no more reliability reports in recent than in earlier years.

Secondly, the reports on reliability which do appear are uniformly high. Review of some 30 studies and experiments reveals a range of correlation coefficients between about .78 and .99 with a concentration at about .90, and a range of percentage agreements between 66% and 96% with a concentration over 90%. However, such magnitudes do not guarantee the reliability of content analysis. Many attempts at content analysis may have been abandoned because of low reliability in the preliminary results. In addition, the published content analysis studies which do *not* report on reliability may be presumed to have had less satisfactory results, or perhaps not to have included reliability tests at all. In short, the high reliabilities which are reported are more in the nature of encouraging indications than absolute guarantees.

Third, most of the reported reliability results apply to relatively simple versions of content analysis. The reliability of a list of symbols (e.g., Grey, Kaplan, and Lasswell, 1949) may be high, but the reliability of a complex semantic analysis is another matter. Presumably the use of complicated and sophisticated categories creates serious problems in reliability. This is the problem of the balance between reliability of the procedures on the one hand and the richness of the categories on the other. What does it matter that we gain reliability if in the process we lose all our insights? This critical area of content analysis, i.e., the reliability of complex categories, still needs to be adequately handled.

Fourth, there is a reasonable likelihood of the *under*estimation of reliability in many studies, arising from the measurement of reliability on detailed categories which are later subsumed into more general categories for reporting purposes. This procedure, as a matter of fact, is recommended by one of the experimental studies of reliability (Kaplan and Goldsen, 1949). The distinction is made between "gross reliability," applying to *all* the coded categories, and "net reliability," applying only to those categories, often cumulations, used in the report (Reiss, 1949).

Fifth, there are differentials in the reliability of various kinds of analysis. Reliability is higher under these conditions: the simpler the categories and the unit, the more experienced and better trained the coders, the more precise and complete the set of coding rules, the fuller the illustrations. In one case, agreements between coders ranged from 67% on inclusion of items to 90% and higher for direction (Allport and Faden, 1940). In another, coder instruction increased agreement from 10-15% (Kaplan and Goldsen, 1949). In still another, an experiment designed to test the greater efficiency of a classification scheme in which the analyst performed only "one psychological operation" at a time, by making a step-by-step *series* of decisions on a complex category, actually showed that this method was a little less reliable than the traditional process of deciding at once whether the content fitted the (combined) category (Schutz, 1950).

In sum, the problem of reliability, on which a good deal in content analysis depends, has been successfully handled in a number of particular studies, but it is by no means a "solved" problem for the field as a whole.

THE PROBLEM OF SAMPLING

The problem of sampling in content analysis is complicated by the fact that as many as three decisions about the sample are necessary for many studies. There are three universes from which samples must be drawn, and in any particular problem the methods of sampling can differ among the three. First, there is the problem of sampling *titles* (i.e., specific newspapers or radio stations or magazines); then there is the problem of sampling the *issues* or *dates* of the titles; and finally, there is the problem of sampling the relevant *content* within specific issues of specific titles.

Titles. As has been indicated, several studies have relied upon samples composed of the largest or the most readily available channels or media units. However, in some studies samples have been selected on the basis of various correlates relevant to the problem under analysis. On occasion the selection is made in terms of *geographical area.* Newspapers from different regions of the country were represented in a study of the editorial treatment of education (Foster, 1938); country weekly newspapers in Minnesota were chosen on the basis of their distance from the Twin Cities (Taeuber, 1932); for a study of foreign language broadcasts listening posts were set up "throughout the United States" (Arnheim and Bayne, 1941). The selection may be made in terms of *types (or presumed audience).* Four types of magazines (general weekly, general monthly, women's, confessionals) were included in an analysis of short stories (Berelson and Salter, 1946). Other studies have used as the criterion for inclusion *editorial position.* In a study of local treatment of press association dispatches on foreign news, "isolationist" and "interventionist" newspapers were selected (Rogers, 1941). Other studies have selected their sample in terms of *size or "importance."* In the study of editorials on education, newspapers were also selected by circulation (Foster, 1938); in the study of country weekly newspapers the earliest one established in each zone was taken (Taeuber, 1932); in the study of radio programming, stations of different power and network affiliations were selected (Baker, 1949). Another criterion that has been used is *time of issue.* In a study of news breaks and news repetition, morning and evening newspapers were included (Buchanan, 1942).

This brief list contains the major correlates by which samples of titles have been selected in content analysis studies. It serves to illustrate the first sampling problem which must be met, namely, which newspapers or magazines or radio stations (or other units) are to be selected.

Issues. Once the titles have been selected, the problem arises of how many issues of the title to include. Now this matter of sampling the title through time should not, of course, be identified with the matter of defining the period under study. That is, the period for analysis is selected as part of the research design, but the problem of selecting issues of the titles within the period (or periods) still remains. For example, an analysis of the treatment of minority groups in magazine fiction was done for the years 1937 and 1943 in order to test the effect of the war, but the problem of selecting specific issues of the designated magazines within those years still had to be faced (Berelson and Salter, 1946).

The problem of securing representativeness in a sample of issues (or dates) is conditioned by certain factors. For example, the relevant content may be subject to seasonal variations, and if the problem necessitates representation of the various periods of the year, samples must be chosen with this in mind (e.g., Berelson and Salter, 1946; Taeuber, 1932; Willey, 1926). At least one study went completely amiss because of inattention to this detail: an analysis of the weddings reported in the society pages of the Sunday *New York Times* in June found no Jewish marriages (Hatch and Hatch, 1947). A critic pointed out that ". . . Jewish marriages are not performed in the seven weeks between Passover and the Feast of Weeks and in the three weeks preceding the day of mourning for the destruction of the Holy Temple in Jerusalem. Almost invariably, June falls either into the one or the other period" (Cahnman, 1948, p. 97). Again there is the problem of what constitutes "average" or "normal" times for certain sample purposes. Since many analyses have demonstrated the crucial importance of events in determining communication content, it has often seemed necessary to "control" this factor. Thus, an experiment devised to test reliability included material from three periods — one of which was characterized by favorable events, one by unfavorable, and one by neutral (Grey, Kaplan, and Lasswell, 1949). Another study was based entirely upon a series of events involving the USSR in both positive and negative contexts — e.g., Soviet aid to Loyalist Spain and the Soviet treason trials, 1937; the battle for Stalingrad and the Second Front discussion, 1942, etc. (Kriesberg, 1947). The most intensive experimental investigation of sampling methods in content analysis dealt with the selection of issues of *Pravda* over a six-month period in 1941 (Mintz, 1949). The general conclusions were that weekly samples were only as efficient as three-day samples for the month and that every-fifth-day and every-second-day samples were "distinctly better."

Content. And even after the titles and the issues have been selected, there remains the problem of which sections or how much of the relevant content to include. After certain issues of certain magazines have been chosen for analysis, it may still be possible to satisfy the study's objectives by taking, say, only half the stories. Or, once certain newspapers on certain dates have been selected, the analysis may be restricted to a particular part of the paper, such as the front page, rather than extended to the paper as a whole. Sampling techniques are applicable on this level of content selection as well.

The last illustration raises the question of whether particular sections of communication channels are representative of the whole. We know that the front page is not representative of the "news" in the entire newspaper. Is it necessary in sampling the relevant content to take into account certain physical factors which can affect the distribution of important categories? Each of the major channels has its own prominent devices, as follows:

Newspaper	*Radio*	*Film*
Space	Length	Length
Location	Order	Order
Page placement	Intonation	Color
Size of headline type	Sound	Intonation
Type face	Music	Sound
Use of color		Music
Use of illustrations		

Such prominence devices may be correlated with relevant characteristics of the content. In fact, the most systematic study of this problem has shown that they are (Stewart, 1943).

In sum, then, sampling procedures in content analysis must often be applied to three different universes: the titles, the issues of the titles, and the relevant content of the issues of the titles. And, of course, the same sampling methods need not be applied in all three cases. In fact, in most cases it will be appropriate to design different samples and use different sampling methods for the three universes. In some cases — and this has been true in several actual studies — it is appropriate to sample only one or two of the levels and take the universe for the other (s). This procedure has resulted in quite different, but still suitable, research designs.

In a few cases different samples have been drawn for different purposes in the same study. This procedure is particularly applicable to studies which investigate problems of both an extensive and an intensive nature. Thus a study of German broadcasts to North America included one sample for general analysis over a seven-month period and in addition took *every* reference to a specific event for two weeks within that period (Bruner, 1941). The latter body of content was subjected to much more intensive analysis than the former.

THE PROBLEM OF INFERENCE

What content analysis as such does is to provide a more or less precise description of the content in terms meaningful for the problem at hand. Yet some content analysis studies present conclusions and interpretations of wider application than the content itself. In the large, then, there are two points at which a content study can go wrong. First, it can go wrong in applying the techniques of analysis, in which case the description of the content itself is mistaken, and so is anything which follows from it. Secondly, it can go wrong in applying the techniques of interpretation or inference to the correct content description. Some studies are open only to the first source of error, some to both sources. This raises the problem of inference in content analysis.

In a great many studies there is no real problem of inference at all. This is true for all those content analyses in which the description of content itself is the primary objective. Such studies can be said to contain implicit inferences about the causes or the consequences of the content — and some contain them explicitly — but such inferences are in the nature of addenda to or reformulations of the basic data. Thus a trend study of newspaper content can be considered material for inferences about the changing character of press controls and/or about the changing character of public attention. But such "inferences" are usually nothing but reformulations in other terms of the result of the content analysis itself. Or take the case of the treatment of ethnic groups in magazine fiction: after the basic content data were described at some length, a few hypotheses about the probable reasons for and effects of the content were proposed, but these were appended to the main job and were clearly presented as such; they could be eliminated without doing violence to the substance of the study. In short, where the central concern is with the content as such — with the content as the category, not the indicator — then the problem of inference is virtually nonexistent.

Nevertheless, it is useful to inquire briefly into the types of generalizations which characterize such studies. Mainly of a historical or descriptive nature, they base their conclusions upon these four kinds of relationships:

Trend comparisons. These studies contain conclusions about changes in the frequency of the categories through time, many of which have already been described. Some studies simply present the changes continuously over a long period of time (e.g., Taeuber, 1932); others present them at two selected points in time (e.g., Lowenthal, 1943); and still others present them for periods of different lengths which have themselves been selected with reference to important external criteria (e.g., Leites and Pool, 1949). In all such cases, however, the description of content along a time continuum provides the basis for the conclusions.

Intracontent comparisons. Here are the studies whose major conclusions rest upon relationships between two or more characteristics or parts of the same body of content. For example, the comparative treatment of majority and minority groups in magazine fiction on a range of selected factors is the basis for generalization (Berelson and Salter, 1946). The comparative emphases in popular jokes dealing with Jewish, Negro, and Irish characters reveal the stereotypes operating in such material (Barron, 1950). The technical comparison of the treatment of Negro-white relations in different parts of Negro newspapers is another case in point (Stewart, 1943). In short, the conclusions here rest upon comparisons *within* the same body of content.

Comparisons of different bodies of content. This is the related application in which conclusions rest upon comparisons between or among different bodies of content. Comparisons of the content of Swiss and British newspapers (Guttinger, 1939); of the treatment of a number of wars in American and enemy textbooks (Walworth, 1940); of the correspondence in propaganda output by two national radios (Berelson and de Grazia, 1947); of the motion picture versions of "good" books with the originals (Asheim, 1949); of the sensationalism in radio and newspaper treatments of the same event (Gottlieb, 1947) — these are examples of this form of generalization. Ordinarily it requires the analysis of the different contents by the same set of categories.

Comparison of content with a standard. In this application the relation between the content and the standard which it is supposed to meet serves as a basis for the conclusion. The standard may be implicit and assumed or it may be explicit and established. For example, in analyses of newspaper and radio treatment of controversial issues, the conclusions are based upon the nonconformity of the content to a presumed standard of "balance" or "fairness" (e.g., Sussman, 1945; Gabel and Gabel, 1946a and b). Prior work in establishing criteria for readability levels, translated into content terms, is applied to subsequent analyses in this field. A considerable number of content analyses base their conclusions upon this form of relationship, although it is not often completely stated in these terms.

A good deal of loose interpretation appears in content analyses because the exact nature of the inference is hidden in a cloud of ambiguous terms, which often turn out to be the verbs which bind the content subject to the noncontent predicate. Typical terms are "reflects," "relates," "expresses," "indicates," "projects," "suggests," "means," etc. To the extent that such terms have value, it lies in their assertion that the two sets of factors have something to do with each other; their deficiency lies in the vagueness of the assertion. *How,* concretely, does the German *Zeitgeist,* whatever it is, come to determine the content of the German film — through the conscious, deliberate exploitation of the public's attitude by discerning producers? Through the unconscious expression of public feelings by the producers, as members of the society? Through their conscious expression of such feelings? What part of the public? What group of movie makers?

On occasion it will be impossible to base at least parts of the inference upon a previously established generalization relevant to the problem. For example, the British predictions about the German possession of V-2 weapons were based upon the historically correct proposition that the Nazi propagandists did not arouse popular expectations of such action without real reason. A similar proposition served as the basis for the American analyst's inference that Goebbels' threat of "counter terror" meant that the Nazis were considering renewed anti-Jewish measures. Where such supports for the chain of inference exist they are invaluable.

Basically it might not be far wrong to suggest that there are two types of inferences drawn from content data — the direct and the inverse. In the direct method, essentially, the content is taken at face value. What it says is, in effect, translated into an inference about noncontent matters. If Fritzsche says that the German people should not expect continued successes from submarine warfare, the inference is that the Nazi elite does not expect such successes. If popular magazines treat subject *A* with high frequency and subject *B* with low, the inference is that people are more interested in the former topic.

In the inverse method, the opposite of the content statement is taken as the inference. If the Nazi radio insists that the Italians are full of fight, the inference is that the Nazi elite thinks that the German people do *not* think well about the sticking power of their allies. If the American movies depict a situation in which the hero is loved unconditionally by the heroine while he remains doubtful in his affection, the inference becomes a "reversal of the experience of American boys with their exacting mothers . . . In this case the relation between a movie theme and one of its probable real life sources is a reversal rather than a reproduction" (Wolfenstein and Leites, 1950, p. 306).

The basic problem, then, is to decide which inference to make — the direct or the inverse. Little progress has been made toward differentiating the conditions under which the one or the other is most appropriate. Suppose an analyst in 2051 notices that American novels of 1951 contain a great many sexual references. What is he to infer about sexual mores in this country — that they were free and open (as revealed by the attention given them in literature) or that they were repressed and inhibited (as revealed by the same)?

In sum, the problem of inference in content analysis is no different from similar problems in other areas of social research. Ultimately the inference itself must be tested by methods directly applicable to it. To that end, full and unambiguous statements of the detailed inference itself, of the line of reasoning that went into it, and of the assumptions contained therein are desirable.

CONCLUSION

This, then, is content analysis. Not all the studies reported or described in this review were well conceived or well executed. Some of them are quite pretentious, some obvious, some pedestrian, some clever but poorly done. No brief is made here for all the studies used for illustrative purposes; in any case, this is what content analysis currently amounts to.

How the contributions of content analysis measure up against those of other methods in social science research is difficult if not impossible to say. Because of the relative availability of the raw material, many sins have been committed in the name of content analysis simply because something else was harder or more inconvenient to do. At the same time, however, the method has produced a number of useful studies which combine original ideas with sound, careful documentation. This is, of course, the heart of the matter.

In content analysis, as anywhere else in social research, it is important to start in the right way. Simply going on a fishing expedition through some common communication material is almost certain to be unrewarding. Unless there is a sensible, or clever, or sound, or revealing, or unusual, or important notion underlying the analysis, it is not worth going through the rigor of the procedure, especially when it is so arduous and so costly of effort. Content analysis, as a method, has no magical qualities — you rarely get out of it more than you put in, and sometimes you get less. In the last analysis, there is no substitute for a good idea.

References

Abt, L. E., & Bellak, L. (Eds.) *Projective psychology.* New York: Knopf, 1950.
Albig, W. The content of radio programs – 1925-1935. *Soc. Forces,* 1938, *16,* 338-349.
Allport, G. W. The use of personal documents in psychological science. *Soc. Sci. Res. Council Bull.,* 1942, *49.*
Allport, G. W., & Bruner, J. S. Fifty years of change in American psychology. *Psychol. Bull.,* 1940, *37,* 757-776.
Allport, G. W., & Faden, Janet M. The psychology of newspapers: Five tentative laws. *Publ. Opin. Quart.,* 1940, *4,* 687-703.
Anderson, H. H., & Anderson, Gladys L. *An introduction to projective techniques.* New York: Prentice-Hall, 1951.
Andrews, T. G., & Muhlhan, Gertrude Analysis of congruent idea patterns as a study in personality. *Charact. Pers.,* 1943, *12,* 101-110.
Arnheim, R. The world of the daytime serial. In P. F. Lazarsfeld & F. N. Stanton (Eds.), *Radio research: 1942-1943.* New York: Duell, Sloan & Pearce, 1943. Pp. 34-107.
Arnheim, R., & Bayne, Martha C. Foreign language broadcasts over local American stations. In P. F. Lazarsfeld & F. N. Stanton (Eds.), *Radio research: 1941.* New York: Duell, Sloan & Pearce, 1941. Pp. 3-64.
Ash, P. The periodical press and the Taft-Hartley Act. *Publ. Opin. Quart.,* 1948, *12,* 266-271.
Asheim, L. From book to film: A comparative analysis of the content of selected novels and the motion pictures based upon them. Unpublished doctor's dissertation, Univ. of Chicago, 1949.

Baker, K. An analysis of radio's programming. In P. F. Lazarsfeld & F. N. Stanton (Eds.), *Communications research: 1948-1949.* New York: Harper, 1949. Pp. 51-72.
Baldwin, A. L. Personal structure analysis: A statistical method for investigating the single personality. *J. abnorm. soc. Psychol.,* 1942, *37,* 163-183.
Bales, R. F. *Interaction process analysis. A method for the study of small groups.* Cambridge, Mass.: Addison-Wesley, 1950.
Barron, M. L. A content analysis of intergroup humor. *Amer. sociol. Rev.,* 1950, *15,* 88-94.
Bartlett, K. G. Trends in radio programs. *Ann. Amer. Acad. polit. soc. Sci.,* 1941, *213,* 15-25.
Berelson, B. The effects of print upon public opinion. In D. Waples (Ed.), *Print, radio, and film in a democracy.* Chicago: Univer. of Chicago Press, 1942. Pp. 41-65.
Berelson, B. The quantitative analysis of case records: An experimental study. *Psychiatry,* 1947, *10,* 395-403.
Berelson, B., & de Grazia, S. Detecting collaboration in propaganda. *Publ. Opin. Quart.,* 1947, *11,* 244-253.
Berelson, B., & Salter, Patricia. Majority and minority Americans: An analysis of magazine fiction. *Publ. Opin. Quart.,* 1946, *10,* 168-190.
Bettelheim, B., & Janowitz, M. *Dynamics of prejudice. A psychological and sociological study of veterans.* New York: Harper, 1950.
Blythe, Irene T. The textbooks and the new discoveries, emphases, and viewpoints in American history. *Historical Outlook,* 1932, *23,* 395-402.
Bruner, J. S. The dimensions of propaganda: German shortwave broadcasts to America. *J. abnorm. soc. Psychol.,* 1941, *36,* 311-337.
Buchanan, R. E. News breaks and news repetition in morning-evening newspaper cities. *Journalism Quart.,* 1942, *19,* 356-361.

Cahnman, W. A note on marriage announcements in the *New York Times. Amer. sociol. Rev.,* 1948, *13,* 96-97.
Child, I. L., Potter, Elmer H., & Levine, Estelle M. Children's textbooks and personality development: An exploration in the social psychology of education. *Psychol. Monogr.,* 1946, *60,* No. 3.

Dale, E. *The content of motion pictures.* New York: Macmillan, 1935.
Dale, E., & Chall, Jeanne A formula for predicting readability. *Educ. Res. Bull.,* 1948a, *27,* 11-20.
Dale, E., & Chall, Jeanne A formula for predicting readability. *Educ. Res. Bull.,* 1948b, *27,* 37-54.
Deutsch, F. Artistic expression and neurotic illness. 1. The respiratory neurosis of Charles Kingsley. *The water babies* and *Anton Locke's* dream. *Amer. Imago,* 1947, *4,* 64-102.
Dollard, J., & Mowrer, O. H. A method of measuring tension in written documents. *J. abnorm. soc. Psychol.,* 1947, *42,* 1-32.

Estensen, E. V. McGuffey – A statistical analysis. *J. educ. Res.,* 1946, *39,* 445-457.

Flesch, R. *Marks of a readable style.* New York: Teachers College, Columbia Univer., 1943.
Flesch, R. *How to test readability.* New York: Harper, 1951.
Foster, C. R. *Editorial treatment of education in the American press.* Cambridge, Mass.: Harvard Univer. Press, 1938.
Foster, H. S. Charting America's news of the World War. *Foreign Affairs,* 1937, *15,* 311-319.
Fries, C. *American English grammar.* New York: Appleton-Century, 1940.

Gabel, M., & Gabel, Hortense Texas newspaper opinion: I. *Publ. Opin. Quart.,* 1946a, *10,* 57-70.
Gabel, M., & Gabel, Hortense Texas newspaper opinion: II. *Publ. Opin. Quart.,* 1946b, *10,* 201-215.

Gottlieb, Lillian Radio and newspaper reports of the Heirens murder case. *Journalism Quart.*, 1947, *24*, 97-108.

Gray, W. S., & Leary, B. E. *What makes a book readable; With special reference to adults of limited reading ability: An initial study.* Chicago: Univer. of Chicago Press, 1935.

Grey, A., Kaplan, D., & Lasswell, H. Recording and context units: Four ways of coding editorial content. In H. D. Lasswell & N. Leites (Eds.), *Language of politics.* New York: Stewart, 1949. Pp. 113-126.

Guttinger, F. Distribution of space in British and Swiss newspapers. *Polit. Quart.*, 1939, *10*, 428-441.

Hale, W. B. *The story of a style.* New York: Huebsch, 1920.

Hamilton, T. Social optimism and pessimism in American Protestantism. *Publ. Opin. Quart.*, 1942, *6*, 280-283.

Harris, H., & Lewis, P. M. The press, public behavior and public opinion. *Publ. Opin. Quart.*, 1948, *12*, 220-226.

Hart, A. *Shakespeare and the homilies.* Melbourne, Australia: Melbourne Univer. Press, 1934. (Ch. 4)

Hart, H. Changing social attitudes and interests. In *Recent social trends in the United States.* Vol. 1. New York: McGraw-Hill, 1933. Pp. 382-442.

Hatch, D. L., & Hatch, Mary A. Criteria of social status as derived from marriage announcements in the *New York Times. Amer. sociol. Rev.*, 1947, *12*, 396-403.

Hayworth, D. An analysis of speeches in the presidential campaigns from 1884-1920. *Quart. J. Speech*, 1930, *16*, 35-42.

Herma, H., Kris, E., & Shor, J. Freud's theory of the dream in American textbooks. *J. abnorm. soc. Psychol.*, 1943, *38*, 319-334.

Jacob, P. E. Atrocity propaganda. In H. L. Childs & J. B. Whitton (Eds.), *Propaganda by short wave.* Princeton: Princeton Univer. Press, 1942. Pp. 211-259.

Janis, I. L. Meaning and the study of symbolic behavior. *Psychiatry*, 1943, *6*, 425-439.

Jones, Dorothy B. Quantitative analysis of motion picture content. *Publ. Opin. Quart.*, 1942, *6*, 411-428.

Jones, E. The problem of Hamlet and the Oedipus complex. (Introduction to Vision Press Edition of *Hamlet.*) London: Vision Press, 1947.

Kanzer, M. Dostoyevsky's matricidal impulses. *Psychoanal. Rev.*, 1948, *35*, 115-125.

Kaplan, A. Content analysis and the theory of signs. *Phil. Sci.*, 1943, *10*, 230-247.

Kaplan, A., & Goldsen, J. *Reliability of certain categories for classifying newspaper headlines.* Library of Congress, Experimental Division for Study of War-Time Communications, Document No. 40, Washington, D. C., 1943.

Kaplan, A., & Goldsen, J. The reliability of content analysis categories. In H. D. Lasswell & N. Leites (Eds.), *Language of politics.* New York: Stewart, 1949. Pp. 83-112.

Kearl, B. A closer look at readability formulas. *Journalism Quart.*, 1948, *25*, 344-348.

Kecskemeti, P., & Leites, N. Some psychological hypotheses on Nazi Germany. *J. soc. Psychol.*, 1947, *26*, 141-183.

Kecskemeti, P., & Leites, N. Some psychological hypotheses on Nazi Germany. *J. soc. Psychol.*, 1948a, *27*, 91-117.

Kecskemeti, P., & Leites, N. Some psychological hypotheses on Nazi Germany. *J. soc. Psychol.*, 1948b, *27*, 141-164.

Kingsbury, Susan M., Hart, H., & associates. *Newspapers and the news: An objective measurement of ethical and unethical behavior by representative newspapers.* New York: Putnam's Sons, 1937.

Kracauer, S. *From Caligari to Hitler.* Princeton: Princeton Univer. Press, 1947.

Kriesberg, M. Soviet news in the *New York Times. Publ. Opin. Quart.*, 1947, *10*, 540-564.

Kris, E. Prince Hal's conflict. *Psychoanal. Quart.*, 1948, *17*, 487-506.

Kris, E., & Leites, N. Trends in twentieth century propaganda. In G. Roheim (Ed.), *Psychoanalysis and the social sciences.* Vol. I. New York: International Univer. Press, 1947. Pp. 393-409.

Kris, E., & Speier, H. *German radio propaganda.* New York: Oxford University Press, 1944.

Landis, M. H., & Burt, H. E. A study of conversations. *J. comp. Psychol.*, 1924, *4*, 81-89.

Lasswell, H. D. *Propaganda technique in the World War.* New York: Knopf, 1927.

Lasswell, H. D. A provisional classification of symbol data. *Psychiatry*, 1938, *1*, 197-204.

Lasswell, H. D. World attention survey. *Publ. Opin. Quart.*, 1941, *5*, 456-462.

Lasswell, H. D. Communications research and politics. In D. Waples (Ed.), *Print, radio, and film in a democracy.* Chicago: Univer. Chicago Press, 1942a. Pp. 101-117.

Lasswell, H. D. The politically significant content of the press: Coding procedures. *Journalism Quart.*, 1942b, *19*, 12-23.

Lasswell, H. D. *Analyzing the content of mass communication: A brief introduction.* Library of Congress, Experimental Division for Study of War-Time Communications, Document no. 11, Washington, D. C., 1942c.

Lasswell, H. D. Describing the contents of communications. In B. L. Smith, H. D. Lasswell, & R. D. Casey (Eds.), *Propaganda, communication, and public opinion.* Princeton: Princeton Univer. Press, 1946. Pp. 74-94.

Lasswell, H. D. Detection: Propaganda detection and the courts. In H. D. Lasswell & N. Leites (Eds.), *Language of politics.* New York: Stewart, 1949. Pp. 173-232.

Lasswell, H. D., & Blumenstock, Dorothy *World revolutionary propaganda.* New York: Knopf, 1939.

Lasswell, H. D., Leites, N., & associates. *Language of politics. Studies in quantitative semantics.* New York: Stewart, 1949.

Lazarsfeld, P. F. The controversy over detailed interviews — an offer for negotiation. *Publ. Opin. Quart.,* 1944, *8,* 38-60.

Lee, A. McC., & Lee, Elizabeth B. *The fine art of propaganda: A study of Father Coughlin's speeches.* New York: Harcourt, Brace, 1939.

Leites, N., & Pool, I. de S. *On content analysis.* Library of Congress, Experimental Division for Study of War-Time Communications, Document no. 26, Washington, D. C., 1942.

Leites, N., & Pool, I. de S. Interaction: The response of Communist propaganda to frustration. In H. D. Lasswell & N. Leites (Eds.), *Language of politics.* New York: Stewart, 1949. Pp. 334-381.

Leites, N., & others. Politburo images of Stalin. *World Polit.,* 1951, *3,* 317-339.

Lerner, D. The American soldier and the public. In R. Merton & P. Lazarsfeld (Eds.), *Continuities in social research.* Glencoe, Ill.: Free Press, 1950. Pp. 212-251.

Lewin, H. S. Hitler youth and the Boy Scouts of America: A comparison of aims. *Hum. Relat.,* 1947, *1,* 206-227.

Lindner, R. N. The content analysis of the Rorschach protocol. In L. E. Abt & L. Bellak (Eds.), *Projective psychology.* New York: Knopf, 1950. Pp. 75-90.

Lorge, I. Predicting readability. *Teach. Coll. Rec.,* 1944, *45,* 404-419.

Lowenthal, L. Biographies in popular magazines. In P. F. Lazarsfeld & F. N. Stanton (Eds.), *Radio research, 1942-43.* New York: Duell, Sloan and Pearce, 1943. Pp. 507-548.

Lowenthal, L. The sociology of literature. In W. Schramm (Ed.), *Communications in modern society.* Urbana, Ill.: Univer. of Ill. Press, 1949. Pp. 82-100.

McCurdy, H. G. Literature and personality: Analysis of the novels of D. H. Lawrence. *Charact. Pers.,* 1939a, *8,* 181-203.

McCurdy, H. G. Literature and personality: Analysis of the novels of D. H. Lawrence. *Charact. Pers.,* 1939b, *8,* 311-322.

McCurdy, H. G. A study of the novels of Charlotte and Emily Bronte as an expression of their personalities. *J. Pers.,* 1947, *16,* 109-152.

McGranahan, D. V., & Wayne, I. German and American traits reflectd in popular drama. *Hum. Relat.,* 1948, *1,* 429-455.

Merton, R. K. *Mass persuasion: The social psychology of a war bond drive.* New York: Harper, 1946. (Ch. III.)

Miles, Josephine Some major poetic words. In *Essays and studies by members of the Department of English, University of California.* Berkeley: Univer. Cal. Press, 1943. Pp. 232-239.

Miles, Josephine *The vocabulary of poetry.* Berkeley: Univer. Calif. Press, 1946.

Miles, Josephine *The continuity of poetic language: Studies in English poetry from the 1540's to the 1940's.* Berkeley: Univer. Cal. Press, 1951.

Mintz, A. The feasibility of the use of samples in content analysis. In H. D. Lasswell & N. Leites (Eds.), *Language of politics.* New York: Stewart, 1949. Pp. 127-152.

Ojemann, R. H., & others. A functional analysis of child development material in current newspapers and magazines. *Child Developm.,* 1948, *19,* 76-92.

Palmer, Georgiana P. The tottoi of Aristotle's rhetoric as exemplified in the orators. Unpublished doctor's dissertation, Univer. of Chicago, 1934.

Punke, H. H. Cultural changes and changes in popular literature. *Soc. Forces,* 1937, *15,* 359-370.

Rainoff, T. J. Wave-like fluctuations of creative productivity in the development of west-European physics in the eighteenth and nineteenth centuries. *Isis,* 1929, *12,* 287-307.

Reiss, A. The accuracy, efficiency, and validity of prediction instruments. Unpublished doctor's dissertation, Univer. of Chicago, 1949.

Rickert, Edith *New methods for the study of literature.* Chicago: Univ. of Chicago Press, 1927.

Rogers, W. C. The presentation of the foreign news. Unpublished master's thesis, Univer. of Chicago, 1941.

Royal Commission on the Press, 1947-1949. *Report.* His Majesty's Stationer's Office, 1949. (Especially: Appendix VII: Investigations into the contents of newspapers and their methods of presenting news in the period 1927-1947. Pp. 238-359.)

Runion, H. L. An objective study of the speech style of Woodrow Wilson. *Speech Monogr.,* 1936, *3,* 75-94.

Schutz, W. Theory and methodology of content analysis. Unpublished doctor's dissertation, UCLA, 1950.

Shanas, Ethel The *American Journal of Sociology* through fifty years. *Amer. J. Sociol.,* 1945, *50,* 522-533.

Shneidman, E. S. (Ed.) *Thematic test analysis.* New York: Grune & Stratton, 1951.

Smythe, D., & Horton, D. Analysis of television programs in New York City, January 4-10, 1951. *News-letter of the National Association of Educational Broadcasters,* Jan.-Feb. 1951.

Spurgeon, Carolina F. E. *Shakespeare's imagery, and what it tells us.* New York: MacMillan, 1936.

Stevens, N. E. The fad as a factor in botanical publication. *Sci.,* 1932, *75,* 499-504.

Stewart, M. Importance in content analysis: A validity problem. *Journalism Quart.,* 1943, *20,* 286-293.

Sussman, Leila A. Labor in the radio news; An analysis of content. *Journalism Quart.,* 1945, *22,* 207-214.

TAEUBER, IRENE B. Changes in content and presentation in Minnesota weekly newspapers, 1860-1929. *Journalism Quart.*, 1932, *9*, 281-289.

WALWORTH, A. *School histories at war: A study of the treatment of our wars in the secondary school history books of the United States and in those of its former enemies.* Cambridge, Mass.: Harvard Univer. Press, 1938.

WAPLES, D., & BERELSON, B. What the voters were told: An essay in content analysis. Graduate Library School, Univer. of Chicago, 1941 (mimeographed).

WHITE, R. K. A quantitative analysis of Hitler's speeches. *Psychol. Bull.*, 1942, *39*, 486-487.

WHITE, R. K. Black boy: A value analysis. *J. abnorm. soc. Psychol.*, 1947, *42*, 440-461.

WHITE, R. K. Hitler, Roosevelt and the nature of war propaganda. *J. abnorm. soc. Psychol.*, 1949, *44*, 157-174.

WHITE, R. K. *Value-analysis: The nature and use of the method.* Society for the psychological study of social issues, 1951.

WILLEY, M. M. *The country newspaper: A study of socialization and newspaper content.* Chapel Hill: Univer. of North Carolina Press, 1926.

WOLFENSTEIN, MARTHA, & LEITES, N. *Movies: A psychological study.* Glencoe, Ill.: Free Press, 1950.

WOODWARD, J. L. *Foreign news in American morning newspapers: A study in public opinion.* New York: Columbia Univer. Press, 1930.

WOODWARD, J. L. Quantitative newspaper analysis as a technique of opinion research. *Soc. Forces*, 1934, *12*, 526-537.

WORMHOUDT, A. *The demon lover: A psychoanalytic approach to literature.* New York: Exposition Press, 1949.

WRIGHT, Q., & NELSON, C. J. American attitudes toward Japan and China, 1937-38. *Publ. Opin. Quart.*, 1939, *3*, 46-62.

WYANT, ROWENA, & HERZOG, HERTA Voting via the senate mailbag. *Publ. Opin. Quart.*, 1941a, *5*, 359-382.

WYANT, ROWENA, & HERZOG, HERTA Voting via the senate mailbag. *Publ. Opin. Quart.*, 1941b, *5*, 590-624.

YAKOBSON, S., & LASSWELL, H. D. Trend: May Day slogans in Soviet Russia, 1918-1943. In H. D. Lasswell & N. Leites (Eds.), *Language of politics.* New York: Stewart, 1949. Pp. 233-297.

YULE, G. U. *The statistical study of literary vocabulary.* Cambridge, Eng.: Cambridge Univer. Press, 1941.

Content analysis — A new evidentiary technique. *Univ. of Chicago Law Rev.*, 1948, *15*, 910-925.

CHAPTER 14

The Cross-Cultural Method

JOHN W. M. WHITING
Harvard University

The cross-cultural method utilizes data collected by anthropologists concerning the culture of various peoples throughout the world to test hypotheses concerning human behavior. Some of the hypotheses tested have been derived from theories of cultural evolution, others from theories concerned with the integration of culture, and still others, particularly in recent years, from theories of individual and social psychology.

The cross-cultural method was first used by E. B. Tylor (1889), who presented at the meetings of the Royal Anthropological Institute of Great Britain a paper entitled "On a method of investigating the development of institutions; applied to laws of marriage and descent." It is interesting that Sir Francis Galton, one of the fathers of modern statistics, presided at this meeting. Although the intent of the study was to support Tylor's particular view of cultural evolution, its importance was that in the paper and the discussion which followed most of the basic assumptions of cross-cultural research were touched upon.

It was not until twenty-five years later, however, that the method was again used and, as before, to test evolutionary theory. Hobhouse, Wheeler, and Ginsberg (1915) published a monograph in which the frequency of occurrence of certain social institutions was determined for each stage of economic development from lower hunters to advanced agriculturalists. Essentially similar (but more methodologically and theoretically sophisticated) studies followed after another lapse of nearly twenty-five years with Simmons (1937) and Murdock (1937).

In the last fifteen years, the cross-cultural method has not only become more popular but has changed in its theoretical orientation. It has drawn upon the theory of general behavior science rather than that of cultural evolution. Ford's (1939) study of the frequency of certain types of behavior presumed to be closely related to biological needs was the first to represent this change, even though it was concerned with the development of an appropriate conceptual system and the establishment of cross-cultural norms rather than with the testing of hypotheses. His later work (1945), and with Beach (1951), is also primarily conceptual and normative.

The first study to be published which was concerned with the testing of hypotheses derived from behavior science was that of D. G. Horton (1943). He used the cross-cultural method to investigate the relationship between the drinking of alcoholic beverages and anxiety, and showed that alcohol reduced inhibition in all societies unless special measures were taken to prevent it.

Murdock's *Social Structure,* published in 1949, represents a landmark in cross-cultural research. In this study he used a sample of 250 societies distributed over the world to test hypotheses concerned with the correlation between kinship terminology and other aspects of social structure; particularly, residence rules, the form of marriage, and rules of descent. He draws specifically upon behavior theory to explain the origin and generalization of the incest taboo.

Studies with a behavior science orientation, but of more limited scope than Murdock's, include B. B. Whiting's (1950) study of the relation between sorcery and social control, McClelland and Friedman's (1952) study of the relation between child training practices and achievement motivation, Wright's (1952) study of the relation between child training and aggression as expressed in folktales, and Barry's (1952) study of the relation between child training and certain formal characteristics of decorative art.

The most elaborate use of psychological theory in cross-cultural research is that of Whiting and Child (1953). In this study certain hypotheses relating to the Freudian mechanisms of fixation, displacement, projection, and the development of guilt are tested. The antecedent variables in this study consist of child-training practices, and the consequent variables consist of magical beliefs and practices concerning the cause and cure of illness. Thus, ethnographic data from primitive societies have been used not only to test theories at the level of culture and society but also theories which are derived essentially from individual psychology.

It might be asked at this point why ethnographic material should be drawn upon to test psychological principles. Why not stick to materials gathered in one's own society, where the language and culture are familiar and where more adequate control of the process of data collection is possible?

The advantages of the cross-cultural method are twofold. First, it ensures that one's findings relate to human behavior in general rather than being bound to a single culture, and second, it increases the range of variation of many of the variables.

With respect to the first point, that of generality, one might ask questions such as the following: Is the Oedipus complex a universal phenomenon or is it peculiar to middle-class western Europeans? Does the relation between frustration and aggression hold in societies other than our own?

Since most social psychological studies are done within the framework of western European culture, one can never be certain whether the discovered relationships are valid for all mankind or whether they are an artifact of some limitation or special circumstance of the culture in which they have been discovered. An example will serve to illustrate this point. Sears and Wise (1950) found as a result of a study of a sample of eighty children living in Kansas City that there was a positive relationship between the age of weaning and the degree to which the infant gave indications of emotional disturbance. That is, the later a child was weaned the more disturbance he showed. This relationship reached the 1% level of significance.

One might ask whether this relation between age of weaning and emotional disturbance is a general human trait or whether it is peculiar to the culture of Kansas City. Anthropologists are all too eager to find exceptions to any rule. Thus, Beatrice Blackwood (1935) reports that the Kurtatchi of the Solomon Islands do not wean their children until they are over three years old and that Kurtatchi children show no signs of emotional disturbance. Does this single case disprove the rule? More than two cases, e.g., Kansas City and Kurtatchi, are needed to answer this question.

Fortunately, the data are available. Whiting and Child (1953) in a cross-cultural study of child training and personality collected material, on the same variables as those used by Sears and Wise, from a sample of seventy-five societies distributed over the world. Evidence was available so that judgments could be made for age of weaning and emotional disturbance for thirty-seven societies. The mean score for emotional disturbance on a seven-point scale for the seventeen societies in which the onset of weaning was at two years or less was 3.5, whereas the mean score for the twenty societies in which weaning took place later than two years was 2.8.

It will be seen from this that in this instance the Kurtatchi are not a single exception but are typical of a tendency for late weaning to result in low rather than high emotional disturbance and thus there really seems to be contradiction between the Sears and Wise findings and the cross-cultural evidence. If the methods of measuring these two variables are accepted as reliable and comparable in the two sets of data, one might conclude that different psychological principles were operating in Kansas City than in the rest of the world. Closer examination of the data, however, reveals a much more plausible explanation and illustrates the second advantage of the cross-cultural method, namely, that of increased range.

The effective range of age of weaning for the Kansas City sample was 0-7 months (only five of the eighty cases were later than seven months), whereas for the cross-cultural sample, with the exception of two cases, it was from twelve months to six years. Thus, there is practically no overlap between the two distributions and, as is shown in Table 1, the two findings complement each other to indicate a curvilinear rather than a linear relation between the two variables. Due to lack of range, neither study tells the whole story.

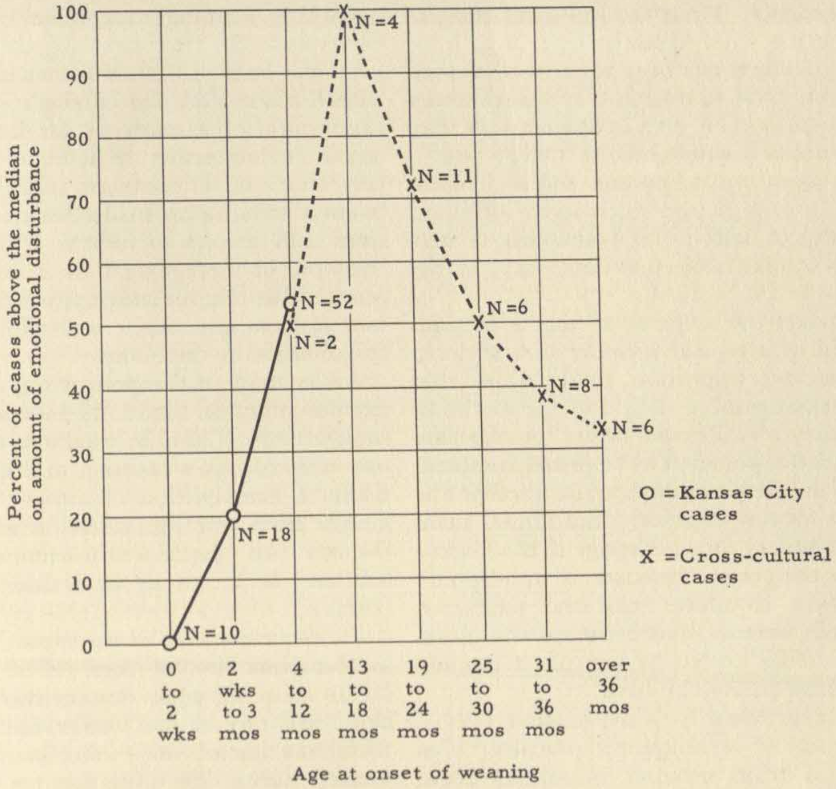
The preceding case illustrates that, with respect to the age of weaning at least, psychologists would be exceedingly hampered by lack of range in developing a theory of its effects which would be applicable the world over. There is considerable evidence that this lack of range is not restricted to the age of weaning. Thus, although much of the accuracy of measurement and experimental control generally found in psychological studies on individuals in our society is lacking in cross-cultural studies, this spread in the variables provides a useful check upon and supplement to psychological research.

Even though extreme cases on some variable may be found in our society, these are deviant from the cultural norms. In cross-cultural research the norm for a whole society is taken as a score. Thus no matter how extreme a case may be from our point of view, the individuals

TABLE 1

RELATION BETWEEN AGE AT ONSET OF WEANING AND AMOUNT OF EMOTIONAL DISTURBANCE SHOWN BY CHILD.

Comparable data from eighty individual children from Kansas City (Sears & Wise, 1950) and from thirty-seven societies (Whiting & Child, 1953) are presented.



who manifest this behavior do not perceive themselves to be deviant. To return to the age of weaning, a Kurtatchi mother who does not wean her child until he is between three and four years old is doing just what other Kurtatchi mothers do. An American mother who nursed her three-year-old child, however, could not help but be aware that she was deviating from the usual practice — a factor which might have considerable effects on her and might therefore contaminate one's findings if one were interested in the consequences of late weaning rather than of deviant behavior.

Furthermore, the cross-cultural method, by studying cultural norms, holds individual variation constant. Psychological studies of individuals in a single society do just the opposite, in that cultural norms are held constant and individual variations are studied. A combination of these two methods should supplement and correct each other in the development of a general theory of human behavior.

Granted that the cross-cultural method is useful in that it provides an increased range for certain variables and a test for the generality of hypotheses, there are some basic assumptions involved in the use of the method which should be considered. First, the method rests on the assumption that the customs of a society are truly comparable to the habits of an individual. There are those who argue that culture is essentially independent of psychological laws. White (1949), for example, holds that there is a science, which he calls culturology, that concerns itself with the evolutionary development, diffusion and transmission of culture, and that the principles which govern these processes are on a different level of abstraction from that of psychology. Another view, represented by Ford (1939), Gillin (1948), Murdock *et al.* (1950), Whiting and Child (1953), and Whiting *et al.* (1953), while recognizing the limitation which social interaction imposes upon culture, has defined a custom (the basic unit of culture) as a special case of a habit. Thus, the psychological principles which apply to habits should, by this definition, apply to customs.

To be more explicit, a custom may be thought to refer to the behavior of a typical individual in a given society. Thus, the Kurtatchi custom of weaning the child between three and four years of age, which has been referred to above, may be understood to refer to a typical or average Kurtatchi mother. Put in another way, the thirty-seven societies presented in Table 1 represent thirty-seven typical parents and their children, one in each of the thirty-seven societies. The problem of defining and measuring a custom in this way is discussed in detail in Whiting *et al.,* Chapter IV.

If one makes the assumption that a custom is the habit of a typical member of a society, then it becomes important to scrutinize the nature of ethnographic data and the methods by which they are collected. Were the societies chosen by anthropologists as large and complex as ours, to describe a custom would involve exceedingly elaborate sampling techniques, such as are described in other chapters of this handbook. It is the general practice of anthropologists, however, to study small and relatively homogeneous societies, or when large and more complex societies are studied, to limit the report to a single village or band.

My own experience with the Kwoma (1941) is not atypical of ethnographic practice. The Kwoma is a tribe situated in central New Guinea. The total population of some nine hundred people are divided into four subtribes. Each subtribe is in turn divided into sibs and each sib into hamlets. Except for a brief trip to the other subtribes and a few visits to other hamlets, virtually all of my observations were made on Rumbima hamlet of the Hayamakwo sib of the Hongwam subtribe of the Kwoma. Rumbima hamlet consisted of thirty-two adults and thirty-two children. Since my house was built in the hamlet and I lived there for nine months, the customs I report are concerned with typical people of various statuses within this hamlet. Whether the customs were representative of the Kwoma as a whole or of the Hongwam subtribe or of only the Hayamakwo sib I cannot say. For cross-cultural purposes, however, Rumbima hamlet — not the Kwoma — should constitute the case.

It is unfortunately the practice, as far as cross-cultural work is concerned, for anthropologists all too frequently to use in the titles of their monographs or articles some group larger than the one which they studied, just as I, myself, did. (*Becoming a Kwoma* rather than "Becoming a Rumbima.") This is misleading, for it suggests that the culture which they describe is representative of the larger group, an assumption which they are generally not warranted in making. Even more important, several ethnographers may each report on a given society but study quite different subgroups whose culture may vary to an unknown degree. If it were the case that one ethnographer described child-training practices in detail and another had a fine collection of folktales it would be very tempting, if one were studying the effect of child training on projective systems, to consider both reports to refer to a single case — tempting but dangerous. It is especially dangerous when the ethnography refers not even to a single society, but to a linguistic stock such as the Sioux or the Zulu.

Fortunately, anthropologists generally identify the group on which they actually gathered their material. This is usually in the introductory chapter. As a caution, to a novice undertaking a cross-cultural study — choose your sample from this identification and not from the title. An excellent discussion of this problem may be found in an article by Murdock (1953).

Although a small homogeneous society such as Rumbima hamlet presents no problem as far as sampling goes, it does raise a question of the stability of one's observation. Thus, in Rumbima hamlet only two children were being weaned during the nine months when I was present. Since both of these children were being observed I obtained a 100% sample of weaning practice at Rumbima in 1937. Two instances, however, do not provide a very stable measure. Individual variation in the two children could obviously introduce considerable bias. In fact, in the present instance, one of the children observed was quite sickly and his response to weaning may have been influenced more by this fact than by the actual age of weaning.

There is, however, an anthropological field technique that compensates to some extent for this defect. Anthropologists generally do not rely alone upon the observation of individuals or the questioning of respondents about their own experience, but they also use informants, who tell them what "the people" do (cf. Chapter 12). Thus, when I asked my Rumbima informants when they (the people of Rumbima) weaned their children, their answer represented not only the two children whom I saw but their opinion based upon all the children which they had observed being weaned. Although this greatly increases the number of individuals observed, it introduces another difficulty — that of the bias of the informant. Needless to say, this is a problem that is always with the social scientist who uses verbal report as a technique for gathering data. To minimize this type of bias, anthropologists generally check all information with a number of informants and they pay particular attention to instances of disagreement. They also check instances of disagreement between the statements of informants and their own, i.e., the anthropologists' observation. As an example of this last point, my Rumbima informants' report about weaning did not correspond to my observation of the sick child. When confronted with this fact, my informants agreed that this was due to the boy's illness.

In general, then, a reasonably reliable estimate of the typical behavior of people in small homogeneous societies or subgroups throughout the world can be made on the basis of ethnographic data.

There is another basic assumption, however, which underlies the cross-cultural method. This assumption is that customs can be compared from one society to another. There are many who argue that every culture represents a complex whole such that a custom which may look the same as a custom in another culture really means something quite different. This same problem is quite similar to that of the psychologist who compares individuals. Child psychologists express the same view when they insist on considering the "whole child."

It is, of course, true that any individual case, be it a person or a society, is unique and, as a whole, incomparable with any other individual case. Scientific investigation, however, is based on the assumption that attributes of the whole can be abstracted and compared. Thus it is said that no two snowflakes are identical, and yet it is also a valid statement that snowflakes are hexagonal, white, and will melt in temperatures above 32° Fahrenheit. Some investigators may be concerned with the beauty and intricacy of the crystalline pattern of each individual snowflake; others with discovering the general truths about snowflakes. Both are legitimate enterprises, although one might not think so if one listened to the heated arguments between proponents of these two approaches.

Those who are concerned with the description, explanation, or diagnosis of a single case generally use rather complex analytic units which combine a very complicated cluster of attributes or variables. Clinical psychologists are wont to speak of the Oedipus complex or the oral syndrome. Similarly, anthropologists in describing a single society speak of the political aspect, the religious system, the social structure, the economic institutions, or the ethos. These abstractions are so complex and multidimensional that they are of little use for cross-cultural research. If no further analysis is made, the ethos of one culture can scarcely be compared with the ethos of another culture. If one views cultures, or individuals, at this complex level of abstraction, one is left with the feeling that they are indeed unique and incomparable.

Comparison demands analysis into simple enough units so that certain attributes may be seen to vary along a single dimension. Comparative anthropologists, whether they have been concerned with historical or functional problems, have tended to use much simpler units of analysis than those who have been more concerned with the description of a single society. For example, Malinowski, concerned chiefly with a complete description of Trobriand culture, used the institution as his basic analytic unit, whereas Kroeber in his early historical comparative work used the concept of culture trait.

The concept of culture trait, although it is more promising than units of greater complexity, has, nevertheless, a weakness as far as testing hypotheses derived from psychological theories. It refers both to behavioral and material events. Cradleboards and methods of child care, spears and hunting techniques, drums and dance forms are all listed as culture traits. Cradleboards, spears, and drums are a different order of events from child care, hunting techniques, and dance forms. Principles of physics and chemistry apply to the former; principles of psychology and sociology to the latter.

The concept of a custom, discussed above, has the advantage of being both psychologically meaningful and simple. It is appropriate here to add to the definition of a custom, which has been defined above as a habit of a typical individual in a society, its essential attributes. They consist of (1) an agent or a subject (2) performing an act (3) in given circumstances. To distinguish a custom from an individual habit it is required that the agent, the act, and the circumstances be generally recognized as categories by the members of the society. Thus, the subject of a custom would be a mother or a farmer rather than Mary Smith or John Jones. Similarly, the circumstances and acts would be recognized categories such as "at dinner time" or "driving a car" rather than events that are relevant only to a particular individual or to some experimental setup. A unit with essentially these attributes was first suggested by Ford (1939), and developed by Murdock *et al.* (1950) and by Whiting and Child (1953).

Not only does the concept of custom have the advantage of simplicity but, since it is defined as a special kind of habit, psychological principles such as those of learning and extinction, generalization and discrimination, and motivation and conflict can be drawn upon to suggest hypotheses about culture. To do this, of course, customs must be measured in the same manner that habits are.

Many anthropologists agree that a culture can be analyzed into its component customs, but are still concerned about the comparison of customs across cultures. They point out that actions which are formally identical have quite different meanings in different societies. Belching may be complimentary or rude; thumbing the nose insulting or meaningless.

This problem is not so difficult as it at first seems. Cross-cultural comparison requires equivalence in meaning rather than formal equivalence. In other words, if customs are defined in terms of universal panhuman dimensions they can be compared across cultures. An insult is an insult whether it be accomplished by thumbing the nose, biting the thumbnail, or shouting "karaganda yikafa" ("little baby" in Kwoma). As Kluckhohn (1953, p. 517) puts it, ". . . biological, psychological, and sociosituational universals afford the possibility of comparison of cultures in terms which are not ethnocentric . . ."

Motives based upon social interaction such as aggression, dependence, affiliation, achievement, and the like provide a set of universal dimensions by which customs that are of particular interest to psychologists can be cross-culturally compared. McClelland and Friedman (1952), Wright (1952), and Whiting and Child (1953) have already used such a device in cross-cultural research.

But how does one go about doing cross-cultural research? What practical problems are involved? The method is not essentially different from any correlational study. The sample must be selected, the variables chosen, scales defined, judges trained, reliability established, care taken to prevent bias in judgments, etc. There are a few special precautions and some useful short-cuts, however, which should be mentioned.

First there is a precaution as to the choice of a sample. A criticism which has been applied to cross-cultural research ever since Tylor (1889) presented his pioneering paper is that of the independence of cases. These critics argue that if two societies derive from a common origin or have extensively borrowed from each other, they should not be counted as two instances but only as one. As Boas (1927, p. 120) stated it:

> In order to make a statistical method a success it is essential that the phenomena counted must be independent of one another. If a number of them go back to the same historical sources they cannot be considered as separate units.

Proponents of cross-cultural research consider this an extreme statement, and take the view that as long as two societies are politically distinct they are as independent as two individuals in a single society. This also is an extreme view. The problem is complex and not as yet completely solved. All empirical evidence so far supports the view, however, that Boas was overcautious. B. B. Whiting (1950) and Whiting and Child (1953) have each used somewhat different methods to correct for this possible error. B. B. Whiting checked her results by recalculating her major correlations on a sample which included only one society from each major culture area in the world. Whiting and Child used the technique of testing some of their hypotheses independently in each of the five continental and oceanic areas of the world.

Even though such corrective devices are available, one should bear the problem of independence of cases in mind in selecting the original sample. A simple precaution is to avoid using two cases which are known to have derived from a common origin within the recent past — that is, within such a short time that their cultures have not had a chance to change. A more cautious rule would require choosing no more than one case from a linguistic area in which the languages were still mutually intelligible. Still another method would be to get a fairly equal representation of cases from the five major areas of the world — North America, South America, Eurasia, Africa, and Oceania.

A second methodological precaution relates to the reliability of sources. The report of a traveler who has visited a society for a few days and writes an account of its culture clearly should not be given as much credence as the report of a trained ethnographer who has lived with the people for a year or more and learned to speak their language. To illustrate the type of report one sometimes finds in travel accounts, there is a persistent story told on the coastal towns of New Guinea by the Australian plantation owners to uncritical visitors. The story runs that there is a tribe of natives living in the interior who have long tails and live in houses on piles. In order to sleep comfortably at night, they have holes bored in the flooring through which they stick their tails. Unfortunately this makes them an easy prey to their headhunting neighbors who creep beneath their dwelling and tie knots in their tails before attacking. Clearly, there is considerable range in the reliability of sources and some assessment of sources should be made before using them.

A third precaution for a novice undertaking cross-cultural research relates to the *identification* of a society. This problem has already ben touched upon above in discussing the problem of sampling. In this context, however, it should be mentioned that the same society may be called by a different name in different sources. Thus the Teton tribe is also referred to as the Western Dakota and the Western Sioux.

A fourth precaution relates to the treatment of societies in which the variables under consideration depend upon a judgment of the presence or absence of certain beliefs or practices. Thus, the belief that animal spirits have the capacity to cause illness might be predicted to occur in societies with a high degree of aggression anxiety. Whiting and Child (1953), in attempting to test this hypothesis, found that this belief was reported to occur in thirty-eight out of their sample of seventy-five societies. In the remaining thirty-seven societies no mention of this belief could be found in the ethnographic sources. In none of these latter societies, however, was there the explicit statement that this belief was absent. This is by no means an unusual state of affairs. If ethnographies consisted of a listing of all the potential beliefs and practices which did *not* occur, as well as those which did, the bulk of the monograph would consist of negative evidence, since any one society has but a small fraction of the total number of customs that may be found throughout the world.

How should societies for which there is no mention of a particular trait be treated? Can it be assumed that the trait is absent if it is not mentioned, or should these societies be omitted on the basis of inadequate information? The latter solution is the more conservative and, as a rule of thumb, should generally be followed if both the antecedent and consequent variables depend upon the presence or absence criterion. Otherwise one may obtain a spurious positive correlation which reflects only differences in the adequacy of ethnographic coverage. That is, both traits would tend not to be reported more frequently in the poorly covered tribes than in the well-covered tribes. Such a spurious correlation will not result when at least one of the variables depends upon some variation in the degree of a trait rather than its presence or absence, and when such variation in degree does not depend upon the amount of information. To assume that no information means absence in such an instance, although it does not introduce bias, does, however, increase the amount of error variance, and for this reason should be avoided if possible.

The practice of always omitting cases with no information, however, often reduces the number of cases so that an adequate test cannot be made. To overcome these difficulties it is suggested, in instances where the presence or absence criterion is used, that a distinction be made between those societies in which no mention is made of a particular custom but the context in which such custom normally occurs is described in detail, and those societies in which the appropriate context is not described. The former can be scored as *inferred absence* whereas the latter can be scored as *not ascertainable* and the case omitted. Thus, for example, if one is interested in "bloodletting" and finds no mention of this practice in an ethnography which has a complete and detailed description of medical theory and therapeutic techniques, it could be fairly safely assumed that the practice was either absent or quite unimportant. If, however, there was no discussion of the medical practices at all, it would be much safer and more appropriate to omit the case.

Fortunately for cross-cultural research, many of the practical problems of selecting a sample, working up a bibliography, and identifying a tribe are greatly facilitated by the Human Relations Area Files. This is a cooperative enterprise among sixteen universities to collect and process the ethnographic material on a worldwide sample of societies. Processing consists of working up a complete bibliography of the major sources on a society, translating those which are in a foreign language, classifying or coding each paragraph according to a standard outline (Murdock, *et al.*, 1950), duplicating the paragraphs on 5 x 8 cards, and distributing them in the appropriate categories in a file. Each member university has a complete set of files. Of particular importance for the problems which have been discussed above is the fact that the files aim at world coverage, with each society carefully identified and each source evaluated. By using these files, a cross-cultural study can be done in a small fraction of the time it would otherwise take.

The selection of cases by the Human Relations Area Files has another advantage for cross-cultural research. The analysis of a sample drawn from these files will be available for the next study based upon the same or an overlapping sample. This is particularly valuable if the analysis includes scalar judgments done with due care for independence and reliability. For the information of those who may wish to take advantage of scores on numerous variables on overlapping samples, they are available in the following sources: Horton (1943), Murdock (1949), B. B. Whiting (1950), Ford and Beach (1951), and Whiting and Child (1953).

The following example will demonstrate the cumulative value of cross-cultural research based on overlapping samples of societies. Ford and Beach (1951, p. 219) report on the duration of the taboo restricting coitus following childbirth. Whiting and Child (1953, pp. 341-343) have a scale measuring the severity of childhood sex training on twenty-nine overlapping societies. On the hunch that the length of the taboo might be related to sex anxiety induced during the child-training process, these two variables were correlated. The results, presented in the following table, show that the societies with high sex anxiety are more likely to have long postpartum taboos than those with low sex anxiety.

Thus far, cross-cultural research based upon available ethnographic literature has been under consideration. There is, however, another method being developed that is of particular interest for psychologists. This method consists of gathering material in a sample of societies for the particular purpose of testing a hypothesis or a set of hypotheses. Such a method obviously cannot be undertaken by a single person. To learn a language, establish rapport, and adapt instruments and then gather material is a long and arduous task. For one person to gather material on enough societies for an adequate sample would take a lifetime or more. It is feasible, therefore, only through cooperative efforts.

An example of a cross-cultural study with the above type of orientation is the values project, "A Comparative Study of Values in Five Cultures," directed by Kluckhohn, Roberts, and Vogt of Harvard University. In this project five southwestern communities in the same environment but with differing cultures are being studied. As a part of the research design of this project, the same variables relating to an overall theory of values have been studied as they manifest themselves in each society.

This type of project not only has the advan-

TABLE 2

RELATION BETWEEN THE DEGREE OF SEX ANXIETY INDUCED DURING SOCIALIZATION AND THE DURATION OF THE POSTPARTUM TABOO RESTRICTING COITUS

	Duration of the post-partum taboo restricting coitus		
	A few weeks to a month	A month to a year	One year to three years
High sex anxiety	Manus Sanpoil Tanala	Chagga Chamorro Riffians	Abipone Azande Chiricahua Dahomeans Dobuans Kwoma Tiv Wogeo
Low sex anxiety	Chenchu Chewa Lamba Lepcha Marshallese Marquesans Murngin Siriono Witoto	Kwakiutl Hopi Thonga	Bena Lesu Trobrianders

tage of increasing the comparability of data gathered from different societies, but it also permits the testing and retesting of hypotheses based upon individual differences in a number of different cultural settings. For example, a number of hypotheses relating to techniques of child rearing and the internalization of moral values were tested on a sample of individual parents and children in each of three of these cultures by Faigin (1952) and Hollenberg (1952). This method bids fair to greatly increase the power of cross-cultural research, since a particular theory can be tested within a society at the same time that an additional case is being added to the cross-cultural sample.

In sum, then, the cross-cultural method, although it is still in its infancy, shows promise of being a useful adjunct to other research methods designed for the development of a general science of human behavior.

References

Aberle, D. F. The psychosocial analysis of a Hopi life-history. *Comp. Psychol. Monogr.*, 1951, *21*, No. 1.

Barry, H., III Influences of socialization on the graphic arts. Unpublished honors thesis, Department of Social Relations, Harvard Univ., 1952.

Blackwood, Beatrice *Both sides of Buka Passage.* Oxford: Clarendon Press, 1935.

Boas, F. Anthropology and statistics. In W. F. Ogburn & A. Goldenweiser (Eds.), *The social sciences.* Boston: Houghton Mifflin, 1927. Pp. 114-120.

Faigin, H. Child rearing in the Rimrock community with special reference to the development of guilt. Unpublished doctoral dissertation, Harvard Univ., 1952.

Ford, C. S. Society, culture, and the human organism. *J. gen. Psychol.*, 1939, *20*, 135-179.

Ford, C. S. A comparative study of human reproduction. *Yale Univ. Publ. Anthrop.*, 1945, No. 32.

Ford, C. S., & Beach, F. A. *Patterns of sexual behavior.* New York: Harper, 1951.

Gillin, J. P. *The ways of men.* New York: Appleton-Century, 1948.

Hobhouse, L. T., Wheeler, G. C., & Ginsberg, M. *The material culture and social institutions of the simpler peoples: An essay in correlation.* London: Chapman & Hall, 1915.

Hollenberg, E. Child training among the Zeepi with special reference to the internalization of moral values. Unpublished doctoral dissertation, Harvard Univ., 1952.

Horton, D. The functions of alcohol in primitive societies: A cross-cultural study. *Quart. J. Stud. Alcohol*, 1943, *4*, 199-320.

Kluckhohn, C. Universal categories of culture. In A. L. Kroeber (Ed.), *Anthropology today.* Chicago: Univ. Chicago Press, 1953. Pp. 507-523.

McClelland, D. C., & Friedman, G. A. A cross-cultural study of the relationship between child-training practices and achievement motivation appearing in folk tales. In G. E. Swanson, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology* (rev. ed.). New York: Holt, 1952. Pp. 243-249.

Murdock, G. P. Correlations of matrilineal and patrilineal institutions. In G. P. Murdock (Ed.), *Studies in the science of society.* New Haven: Yale Univ. Press, 1937. Pp. 445-470.

Murdock, G. P. *Social structure.* New York: Macmillan, 1949.

Murdock, G. P., *et al. Outline of cultural materials: 3rd revised edition.* New Haven: Human Relations Area Files, Inc., 1950.

Murdock, G. P. The processing of anthropological materials. In A. L. Kroeber (Ed.), *Anthropology today.* Chicago: Univ. Chicago Press, 1953. Pp. 476-487.

Sears, R. R., & Wise, G. W. Relation of cup feeding in infancy to thumb-sucking and the oral drive. *Amer. J. Orthopsychiat.*, 1950, *20*, 123-138.

Simmons, L. W. Statistical correlations in the science of society. In G. P. Murdock (Ed.), *Studies in the science of society.* New Haven: Yale Univ. Press, 1937. Pp. 495-517.

Tylor, E. B. On a method of investigating the development of institutions; applied to laws of marriage and descent. *J. anthrop. Inst. Gr. Brit. Irel.*, 1889, *18*, 245-269.

White, L. A. *The science of culture.* New York: Farrar, Straus, 1949.

Whiting, B. B. Paiute sorcery. *Viking Fund Publ. Anthrop.*, 1950, No. 15.

Whiting, J. W. M. *Becoming a Kwoma.* New Haven: Yale Univ. Press, 1941.

Whiting, J. W. M., & Child, I. L. *Child training and personality.* New Haven: Yale Univ. Press, 1953.

Whiting, J. W. M., *et al. Field manual for the cross cultural study of child rearing.* New York: Social Science Research Council, 1953.

Wright, G. O. Projection and displacement: a cross-cultural study of the expression of aggression in myths. Unpublished doctoral dissertation, Harvard Univ., 1952.

CHAPTER 15

The Social Significance of Animal Studies

D. O. HEBB AND W. R. THOMPSON*
McGill University

We do not restrict ourselves in the present chapter to making a review of studies of social behavior in subhuman animals; in fact, not even a thorough review of this kind can be attempted. Our aim instead is to deal with the animal work that may help in understanding human society. In doing so, we must emphasize intellectual processes, and we shall give a large proportion of our space to emotion and motivation in the individual animal. Thus we neglect some important studies of social organization, especially in the ant and the bee (but also in birds and mammals when it does not seem that anything in particular is added to the understanding of man). Our treatment is largely determined by the look of open astonishment that we have often seen at the idea that the study of animals has any interest for social psychologists. We cannot do justice in one chapter to the literature of comparative social psychology, but we do hope to convince the reader that he should read further, and that he need not think it a social error to be found with a copy of the *Journal of Comparative and Physiological Psychology* in hand.

The study of animal behavior is significant in social problems for the same reason as in the problem of learning; we shall, in fact, try to show that social psychology must be dangerously myopic if it restricts itself to the human literature. Our aim is thus to treat here the subhuman-animal data which relate to the problems, methods, and theories of human social psychology. After looking briefly at the logic of this relation, our discussion will concern itself first with intellectual factors in communication and cooperation and will then go on to motivational factors, although we recognize fully that this is an artificial separation. Finally, we shall try to add all this up in a theoretical approach to human society from the comparative point of view.

VALIDITY OF ANIMAL STUDIES

There is a common opinion that understanding human behavior will result from the study of human behavior only, on the ground that proving something true of an animal does not prove it for man. The latter part of this statement is both correct and irrelevant. It implies that an empirical science proceeds by proof and disproof, which is a most misleading idea.

No scientific theory can ever be proved true; nor indeed can we ever be finally sure of having proved one wrong. Too often in the history of science the demolished hypothesis has turned up again full of new life, following some slight modification of its assumptions. Lavoisier's oxygen theory of combustion was early disproved (Conant, 1947); Newton's corpuscular theory of light, once discarded on the assumption that light must be either corpuscular or undulatory, but not both, is now in good standing again in somewhat modified form. As for proving that a hypothesis is true, it is notorious that ten thousand confirmations cannot prevent some spoilsport from thinking tomorrow of a new experiment that upsets it at one blow.

Thus the statement that animal experiment proves nothing for man does not go far enough: finding that something is true *of man* up till now, and for the limited sample of men that one has actually been able to test, cannot logically establish that it will be true of other men in the future, as Hume noted long ago. Finding that a psychotic behaves in a certain way does not prove anything about normal persons, but no one would argue from this that clinical studies have not helped in our understanding of the normal. In principle, study of animals has exactly the same status.

All very fine, says the social psychologist, but when in fact has animal psychology lived up to this billing, and helped in the way that abnormal psychology certainly has? As a first answer we choose an example that may seem to have

---

*We gratefully acknowledge the critical advice received from graduate students at McGill University and from Drs. D. Bindra, D. Krech, A. S. Luchins, H. W. Nissen, T. C. Schneirla, D. C. Williams, and J. P. Zubek.

little bearing on social psychology — partly because it does have a bearing, and partly because it clears up the logic of the way in which animal experiment contributes to a knowledge of man.

STUDIES OF ANIMAL INTELLIGENCE

The example is from the study of individual problem solving in the rat. There is a confused and extensive literature on the relation of environment and heredity to the development of human intelligence. One group of studies has seemed to show that environment can have little to do with intelligence; another that it must have a great deal. In either case the argument is somewhat indirect and not clearly decisive, and the literature itself has not led to a resolution of the apparent contradiction. Indeed, both Goodenough (1946), and Jones (1946), in reviewing the literature, express dissatisfaction with the ambiguities of human experiments on the subject. Studies on the effects of nursery school training on I.Q., for example, do not agree whether there is any increase in measurable intelligence. And the few case studies available on single individuals who have been severely restricted in infancy (Mason, 1942; Hill and Robinson, 1929; Davis, 1940, 1947) usually involve so many uncontrolled variables that interpretation is exceedingly difficult. The situation obviously calls for a systematic human breeding experiment together with a radical variation of the environment during growth. Of course no such experiment can be done. And of course if it is done with animals it would prove nothing for man. However, the animals experiments have been done, and the significance they do have, *vis-à-vis* the human problem, is to show that the entire contradiction may stem from implicit assumptions about man for which there is no factual support. The animal experiments, furthermore, suggest that they are wrong.

Experimental studies have shown clearly that when the postnatal environment is held constant the level of learning capacity and problem solving in the rat is innately determined (Heron, 1935; Tryon, 1940; Thompson and Bindra, 1952); that is, one can breed rats that will be brighter, or duller, when brought up in the same environment. But they show equally clearly that when heredity is held constant between experimental groups, by the split litter method, infant experience determines the level of adult problem solving (Hymovitch, 1952; Bingham and Griffiths, 1952; Forgays and Forgays, 1952). The rat brought up in a "free environment," or even with almost no freedom of bodily movement but with wide and extensive perceptual experience, is the better problem solver at maturity. For the rat, then, adult intelligence depends both on heredity and on the stimulating action of the postnatal environment.

Now the implicit assumption has been made in interpreting the human data that the effect of learning on intelligence can be estimated at any time during growth; if special experiences at the age of 10 or 12 have little effect on the I.Q., if a radical change of environment at age 6 has only a slight effect on the I.Q. by age 7, it is concluded that the kind of experience the child has, at any time during his growth, is not an important factor in his intellectual development. This is, in effect, the assumption that the learning process is the same at all stages of growth, as is also the generality of transfer effects.

But we have no evidence to show that this is so. The animal evidence, furthermore, indicates that the assumption is wrong. If instead we assume for man, as for the rat, that experiences early in development may have a widely generalized and lasting effect which they do not have at a later age, the contradictions of the human data disappear. It is still necessary to see how the principle is to be applied to man in detail. We still must discover at what ages the various aspects of his experience have their lasting effects, and what these experiences are. But finding that the adult I.Q. is unmodifiable, or that putting the underprivileged 6-year-old into a good school does not greatly raise his I.Q. in the next year or so, is no longer inconsistent with the idea that the man reared from birth in an impoverished environment has a lower I.Q. than he would otherwise.

Thus the first point we make in citing this work is that animal experiment may clarify a human problem without "proving" anything. It may draw attention to facets of human behavior one has not noticed; it may point to a troublemaking but implicit assumption; it may suggest a new principle of human behavior. Furthermore, animal experiment in the past has repeatedly shown that the treatment of some human problem or other has been oversimplified. But in all these cases, the relevance of the animal work is strictly dependent on whether, when applied to man, it does clarify the human problem.

The closer one stays to man's behavior the less one can hope to see its broader outlines. One takes for granted too easily the things that all men have in common without seeing how extraordinary some of them are. We concede that animal psychology has had its own myopic errors, but these have usually arisen when study is concentrated on some one laboratory animal, rat or cat, regarded simply as a convenient substitute for man and not as the object of comparative study. For effective recent discussions of the distinction between animal psychology and comparative psychology see Beach (1950) and Schneirla (1952).

THE EARLY ENVIRONMENT

The second point we wish to make, in connection with the experiments cited in the preceding section, has to do with the role of the early environment in the development of social intelligence and social attitudes. The problem is on all fours with the "mental hygiene" problem of emotional development. What childhood experiences make for an ability to "understand people," to be sensitive to another's mood, to know how a contemplated action will affect someone else? We can no longer assume that the insensitive person simply fails to inherit the necessary social intelligence. He may or may not. The social problem of education is not limited to the guidance of emotional development and the establishment of desirable attitudes toward others.

Mental hygiene in this sense is not enough. It would be silly to think that a young woman would have much chance of being emotionally stable if she continually blundered in her human relationships, frustrating herself socially; or that a politician could put "democratic" principles into effect, no matter how good his intentions, without being able to recognize ability and honesty and enthusiasm, or without knowing how to lead except by giving orders. There is no proof whatever that man's social intelligence, his social skills, must be fostered in childhood to reach full development, but everything we know suggests that this is so. For a true mental hygiene, therefore, as well as for a better social order, we must find out in what circumstances the child's social intelligence reaches its full development.

As to social attitudes, again, there is no reason to think that all the recently emphasized work on the role of the early environment in lower animals can be applied directly to man, but there is little doubt that it applies in principle, when we can homologize properly. Hoarding in the adult rat (Hunt, 1941; Hunt, Schlosberg, Solomon, and Stellar, 1947), increased by hunger in infancy, may or may not be homologous with miserliness in man. It is even harder to draw a human moral directly from the very important experiments of Lorenz (1935, 1952) on "imprinting" in young ducks and geese. But taken more broadly, with other studies showing a general importance of early experience on social attitudes and emotionality, such as those of Hall and Whiteman (1951), Bernstein (1952), Levy (1934), Ross, Smith, and Denenberg (1950), and Thompson and Heron (1953), the observations in question assume greater significance — presenting as it were undiluted examples of a kind of lasting fixation that may be fundamental in mammalian behavior but obscured or overlaid by the greater complexities of learning in higher forms. This possibility is brought closer to man by Nissen's (1951) estimate of the relation between sexual deviations in the chimpanzee and extensive early exposure to man.

Failing the possibility of radical rearing experiments with human children, it seems that comparative study offers a solid line of advance in this field. Studies of emotionality and social attitude in a number of species will provide some basis for tentative extrapolation of the curves of phylogenetic development through chimpanzee to man, where the ideas thus arrived at can be tested for their value in clarifying clinical and naturalistic observations. The striking parallel in the implications of von Senden's (1932) data on human congenital blindness and Riesen's (1947, 1950) on chimpanzees reared in darkness, and the equally significant study by Nissen, Chow, and Semmes (1951) on the role of early experience in somesthesis, show that in one area at least, that of perceptual function, some of the effects of the early environment on human behavior could have been predicted from animal experiment. The problem with respect to social attitudes should not be intrinsically more difficult, and the more the points that are determined in the phylogenetic curve the more certain the extrapolations can be.

Social Phenomena among Insects

President Lowell of Harvard is said to have introduced W. M. Wheeler for an honorary degree as a distinguished scientist who had shown that ants, too, can build a complex society without recourse to reason. We would not argue against Lowell that human society is the product of dispassionate deliberation. We do urge that even its primitive forms would be impossible if its members did not have their capacity for thought and speech. Whether he uses his reason in a reasonable way or not, it is quite clear from phyletic comparison that the first key to man's social structure lies in his distinctive intellectual qualities.

The point can be put briefly: the ant's social mechanisms are sensory, man's ideational. This is a distinction that is fundamental to our whole treatment. Accordingly, we shall first clarify this use of terms, and go on in the present section to describe some of the social behavior of insects. It contains striking parallels with human society. But these parallels are superficial, as most of the rest of this chapter will try to show.

On the use of terms: first, though our definition may not be rigorous, we shall consider a social act to be one that is determined by past or present stimulation, or expectancy of stimulation, from another member of the species; or done with the expectancy (intention) of evoking a response from him. This does not wholly exclude interaction between members of different species, if the behavior of the two is at all similar and produces relationships similar to those within a species; a pet dog is sometimes considered, and considers himself to be, a member of the family. But we do exclude pet goldfish; together with cattle-herding, fly-swatting, lion-taming, salmon-fishing, and other forms of predation. It is important also to exclude pseudosocial behavior, in which grouping occurs only because of some stimulus external to the group: examples are animals running from a forest fire, moths attracted to a light, and (a slightly closer approach to social phenomena) situations in which a by-product of one animal's existence attracts another, as in the association of a worm-seeking robin with a ploughman, or the huddling of cattle for warmth.

Next, for an adequate description of what happens in phylogenesis, there is no escape from distinguishing between behavior that is more and less stimulus-bound (Goldstein, 1940) or, more broadly, situation-bound. This is Schneirla's (1946) "biosocial" versus "psychosocial" distinction. The biosocial act is one which is dominated by the present environment (including humoral factors which are part of the environment of the nervous system) as distinct from the psychosocial act, in which "ideation" and "purpose" appear to be crucial.

The empirical criteria of sense-dominated or biosocial behavior are immediacy of response and the possibility of eliciting the same response repeatedly by presentation of the same stimulating situation. The predictability may be for the individual animal only, but the picture is clearest when the same stimulus-response relationship holds for the whole species (which does not necessarily mean that it is entirely "unlearned"). Examples: the response of bird or fish to one of the "releasers" of Lorenz and Tinbergen (Tinbergen, 1951), or the dog's aggressive response to the growl of another dog.

The psychosocial act, under ideational control, is harder to define — there may in fact be several categories here if we know how to distinguish them — but in general it falls under the heading of behavior which has led in the past to use of the term expectancy, or some other reference to a nonsensory determinant of response. The behavior is well adapted to the environment and hence partly under sensory control, but is not predictable from the present stimulation alone and makes it necessary to postulate expectancy, set, attitude, or selective attention as codeterminants of the response.

The reader is likely to have difficulty with this distinction if he regards it as absolute, or primarily due to some theory or other. It is primarily empirical. It is not a distinction between learned and unlearned, and our emphasis is not on supposed differences in the complexity of the response mechanism. One can hardly be certain that all "ideational" responses, about which we know little, are more complex or "higher" than all reflex responses, some of which are known to be complex indeed. Our emphasis instead is on differences in the directness of control of response by sensory and humoral influences.

The most elaborate social structures occur far apart in the phyletic scale: man's at the top, the social insects' near the bottom. In between, organization is loose and sporadic. Apart from the division of labor involved in sex behavior and in parental care, there are few cases in which subhuman vertebrates are essentially dependent on other members of the species for specialized functions that they do not have themselves. Individual animals may act in a way that serves the whole group, but the service could be performed about as well by other members and there seems to be nothing in the behavior of the gregarious birds and mammals that parallels the specialization of function in the castes of social insects, or the specialization of the human officeworker who might die if he were suddenly left without food and clothing produced by others, before he learned how to produce the tools or weapons he would need to support life.

The figures of speech, *caste, queen, soldier,* and *worker,* point up this analogy between social insects and man. But as Schneirla (1946) in particular has insisted, it is an analogy, not to be pushed too far. Despite having only a rudimentary social structure, the chimpanzee has more to tell us concerning the nature of human society than all the social insects taken together. However, the phenomena of the insect's social organization are fascinating in themselves, and (as we have said) they do clarify the human problem by setting it in perspective.

THE PHENOMENA OF INSECT SOCIETIES

The founding of a colony varies only slightly with the family. In most wasps, ants, and termites, the fertilized female leaves with or without a small company of swarming workers to found a new home. In the bumblebee, a queen, alone, selects some cavity in the ground or in a log and at first behaves like a solitary bee. Later, the social behavior characteristic of a worker appears in the care of her young. Still later, when there are workers, she takes the role of the queen only. The progression of the colony from a rather primitive to a more complex social organization is also characteristic of ants. In a new colony the first brood is always composed of workers only. These take care of the second brood, in which soldiers appear.

The division of labor in human societies is approximate and flexible. In the social insects it is much more exact and rigid, based on the needs of reproduction, nutrition, and protection which underlie colony life. Each insect caste is dedicated to a certain function, the queens to reproduction, the workers to nutrition, and, in wasps, to protection also, although in ants and termites this latter function is served by soldier castes. Each caste is highly specialized structurally to perform its appropriate task. In a manner, though there seems little to be gained by it (cf. Schneirla, 1946), they may be considered as analogous to parts of a "supraorganism" (Allee, Emerson, Park, Park, and Schmidt, 1949).

The determination of castes is probably genetic as well as nutritional. In the cases of the two sexes (and thus of the two reproductive castes) it appears to derive from genetic differences in the egg (Light, 1942). But in bees and wasps, determination of queen and worker castes depends largely on nutritive factors. The larvae that are to become queens are fed entirely on "royal jelly," the worker larvae on "royal jelly" only for the first three days and then on "bee nectar." (The workers are a sterile caste which may be said to have been "nutritionally castrated," to use Wheeler's phrase.) The queen reaches maturity in about sixteen days, as compared with twenty-one for workers and twenty-four for drones. Similarly, in ants, caste differentiations between queen, workers, and soldiers of the same sex appears to be determined more by environmental than by hereditary factors. It is known, for example, that ant colonies with a small supply of food will produce only workers (Light, 1942), and again that soldier ants in a colony over a given threshold number will inhibit the development of more soldiers (Gregg, 1942), this possibly being accomplished by the soldiers' exuding an exohormone that passes on to developing individuals or sometimes to the unlaid eggs (Allee *et al.*, 1949). In general, caste determination appears to be fairly flexible in the first stages of growth in the social insects but, once determined, an individual's caste seems to remain rigidly fixed.

The structural differences of the castes, underlying the functional ones, are often striking, involving gross differences in size as well as in specific characters. In one genus of ant (*Carebara*) the queen is several thousand times as large as the worker (Wheeler, 1922). The wasp workers which double as soldiers are equipped with a sting, unlike the drones; but in some ants and termites, the soldier may guard the colony merely by blocking the entrance with his head. In other species, the soldiers are equipped with large, powerful jaws. Many mandibulate termite soldiers and all the nasute have a large frontal gland which excretes a sticky liquid or "latex" capable of disabling enemies.

In addition to the different castes, there are often "guests" and "slaves," more often in ant colonies, but also in those of bees, wasps, and termites. Wasmann (1905), Wheeler (1928), Plath (1935), and others have discussed these phenomena fully. Treatment of guests ranges from bare tolerance to definite encouragement. An interesting example of an ant guest (*Leptothorax emersoni*) is described by Wheeler. It is associated with a larger host (*Myrmica canadensis*) and lives in small chambers near the surface of the latter's nest. The smaller ants have the habit of climbing on the backs of passing host workers and licking their bodies, especially the heads and mouthparts. The hosts permit this and occasionally regurgitate droplets of food which the guests eat. The guests, on the other hand, do not allow *Myrmica* into their own colony, except under special conditions. Many species are able to induce *Phytophthora* (plant lice, scale insects, mealy bugs, etc.), by stroking them with their antennae, to regurgitate honeydew the latter have sucked up from plants. These guests are often housed in specially constructed caverns ("barns") in the nest, and thus were called by Linnaeus "the dairy cattle of the ants."

The extraordinary social phenomena of the insect world include slaves also. Ant slavemakers have been classified into two types, "obligatory" and "facultative," depending upon whether slaves are necessary for existence or not. Four known genera of ants are slavemakers (*Formica, Polyergus, Strongylognathus,* and *Harpagoxenus*). One commonly has slaves but can get along without them, three are of the obligatory class and must have slaves to survive. The slaves are generally taken as larvae in raids on other colonies, hatched and raised in the slavemakers' nest, and put to work excavating the nest, feeding and caring for the young, and doing other chores normally carried out by the workers of that colony. *Polyergus* ants are completely dependent on their slaves and often deplete a whole nest of its larvae and pupae, although the attendant workers are not harmed unless they resist (Plath, 1935). A similar phenomenon, but this time analogous to child labor instead of slavemaking, is found in another species (*Oecophylla smaragdina*): the worker ants use silk produced by their larvae to bind leaves together for the nest. The head of a larva is applied to each leaf in turn by the worker's holding it in the appropriate position. During the process the larva "cooperates" by bending its head whenever it is lowered to a leaf (Plath, 1935). Many other similar examples could be cited.

COMMUNICATION IN INSECT SOCIETIES

One essential key to social organization is the means of communication, in the broad sense of the way in which the behavior of one animal induces cooperative behavior in another. As a prelude to examining the development of communication, let us see what the cues are upon which the complex social organization of insects depends.

For the ant, the "signals" are mainly tactual, chemical, and auditory (Allee *et al.,* 1949). Of these, the chemical may be dominant. An ant with a strange odor is attacked, even if a member of the colony; a foreigner, experimentally provided with the colony odor, can be introduced safely (Dropkin, 1941). But this is by no means representative of all insect interaction. Wheeler (1928) reported that the hungry larvae of a wasp colony protrude their heads from their cells and may actually scratch on the walls of the cell, thereby attracting the attention of wasp nurses which, in feeding the larvae, are rewarded by drops of sweet saliva excreted from the mouths of their charges. A most interesting type of communication occurs in those species of ants and termites whose soldiers block the exits to the nest with their heads. When a worker wishes to leave or enter the nest, she strokes the soldier's head or abdomen with her antennae, whereupon the living "door" moves back to let her pass. After she has gone, he moves back into position (Wheeler, 1928). The ability of the soldier to discriminate a friend from a stranger is, in itself, a remarkable feat. How it is done is not known, nor whether it is learned (that is, whether the discrimination requires practice or improves with repetition).

The most remarkable communication occurs in bees, brilliantly demonstrated in systematic experiments by von Frisch (1950). As he remarks (p. 85), the results at first glance seem fantastic, but they have been attested by Thorpe (1950) and others. The worker who has found a food source is able without leaving the hive to give her fellows not only the direction but also the distance of the food, by the "wagging dance." The farther the food, the fewer the number of turns in the dance. The angle of the dance on the vertical comb, relative to gravity, gives the direction relative to that of the sun. If the sun is hidden, polarized light from an area of blue sky still provides orientation, and it appears that the polarization is the essential cue whether the sun is out or not; no communication is possible if the sky is completely overcast. Through this extraordinary behavior other workers are enabled to find the food source.

LEVELS OF COMMUNICATION AND COOPERATION

Such phenomena raise at once the question of their similarity to human communication. Does von Frisch's evidence mean that the bee has the equivalent of a human sign language, though with a small vocabulary? The question might be phrased, Is the bee's behavior symbolic?, but the term symbolic has been used in so many ways that we find it confusing. A main purpose of this section is to find out how to ask the question better, to find better ways of defining the essential differences of communication at various phyletic levels.

Most insect communication is clearly in the class of the sense-dominated social act. The typical stimulus is a scent, touch, or sound that evokes a typical response, not varying much at different times. The invariability of insect behavior has often been exaggerated, but certainly the relation between stimulus and response is much closer than with mammals.

The behavior described by von Frisch does not appear to be essentially different, although it does appear unusually complex. The wagging dance or "message" involves a delayed response but is directly controlled otherwise by the special effects of polarized light on the complex eye of the bee. The number of turns in the dance may be determined by fatigue or in some other equally elementary way, since it actually corresponds to amount of effort or elapsed time rather than distance: it decreases with distance, and also with a head wind, but increases with a tail wind. No one, of course, would suggest that the bee performs some kind of trigonometric computation which is then expressed in the dance, and remarkable as the whole behavior is, it may still lie in the reflexive class.

Similarly, the problem of the way in which other bees receive the message is unsolved. It has not been shown that the behavior has the characteristic of intentionality (Bierens de Haan, 1929) or purposiveness (Maier and Schneirla, 1935), which is essential before one can conclude that the bee has a language. This point is spelled out in the following section, but we can anticipate here to the extent of suggesting the sort of evidence that might be looked for. Evidence of intentionality might be provided, for example, if only the first of several returning bees made the wagging dance — if the message is already being conveyed to the colony by ten other bees there is little sign of purpose in the behavior that conveys it once more. Or again, if one could remove the returning bee's audience experimentally, would she still make the wagging dance, as if the audience were there? We have no intention of making light of this extraordinary behavior, but understanding it requires that we see its limits as well as its extent. The evolutionary origins of so-called instinctive behavior in general are as difficult to understand, as miraculous, as those of the so-called intelligent behavior of higher mammals; but let us avoid mixing up our miracles.

PURPOSIVE COMMUNICATION AND LANGUAGE

Bierans de Haan and Maier and Schneirla have reserved the term *communication* for purposive communication, the act that is done "with regard to the effect it has on other individuals." We propose here to use the term more broadly but not to lose their essential distinction. By including a class of "reflexive" (nonpurposive) communication, we are able to recognize von Frisch's phenomenon as communication and still not confuse it with phenomena of a psychologically higher level.

The conception of purposive behavior is not, of course, a vitalistic one. The difficulty in identifying behavior as such is that it may have to be seen repeatedly in order to be sure that its objective is achieved neither accidentally nor mechanically: that it shows "means-end-readiness" (Tolman, 1932), or a variation with circumstances that tends toward the same end effect. A bird's "warning" cry would be purposive, for example, if made only when the young were exposed and stopped when they got under cover, or if the warner approached the warned and called quietly, without drawing the marauder's attention also. The essence of the purposive communication is that the sender remains sensitive to the receiver's responses, during sending, and by modification of his sending shows that his behavior is, in fact, guided by the intention (expectancy) of achieving a particular behavioral effect in the receiver. The "broken wing" behavior of the grouse, for example, is purposive, since it is modified according to the marauder's behavior in such a way as to draw him away from the hiding chicks; but the warning cry of the gull, although it tends to produce a protective immobility in the chicks, seems quite reflexive. Carpenter (1942) points out that the alarm calls of a group of primates, startled by an observer, continue long after they have got away out of danger. The calls are therefore more emotional expression than purposive.

Birds and animals do show purposive communication. It seems likely that the crow, for example, or the parrot, may on occasion call or "speak" with intent to influence another. A dog's bark, or a chimpanzee's begging gestures, may be unmistakably purposeful. This does not mean, however, that the bird or subhuman mammal has language. It is important to recognize that purposeful communication occurs in animals, and just as important to recognize that language is something more.

The criterion of language is not only that it is representational, for representational acts occur in animals. Man has what neither chimpanzee nor parrot has: the capacity to combine *and readily recombine* representative or symbolic noises (words), movements (gestures), or modifications of inert things (writing, carving). We propose therefore that the minimal criterion of language, as distinct from other purposive communication, is twofold. First, language combines *two or more* representative gestures or noises purposefully, for a single effect; and secondly, it uses the *same* gestures in different combinations for different effects, changing readily with circumstances. As Morris (1946) and Schneirla (1946) have noted, language has other distinctive characteristics, but we believe that these criteria are enough to set it off fully from animal communication. Too much attention, we believe, has been given to man's special vocal behavior. In sign language and gesture man's distinctiveness is equally clear (Révész, 1944), and we avoid the confusing issue of whether the subhuman primate has the vocal equipment for producing human sounds.

From this point of view, the outstanding fact is man's capacity for a varied combination of symbolic acts, and we therefore propose to call language *syntactic behavior.* We thus have three classes of communication: (a) reflexive (not true communication); (b) purposive but nonsyntactic; and (c) syntactic (and usually purposive), or true language.

There are evidently differences of level within each of these classes, and though it seems clear that true language is unique to man, we can trace an increasing complexity of purposive behavior in subhuman animals which suggests that the discontinuity between ape and man, in communication, is less than it appears. The chimpanzee may well be close to a liminal level that would make language possible (Hayes and Hayes, 1951a).

Underlying syntactic behavior there must be a certain independence or autonomy of conceptual processes. Two or more concepts must be capable of being aroused simultaneously, instead of one crowding the other out; and, as the term *concept* itself implies, they must be arousable by associative mechanisms in the absence of the sensory events that originally gave rise to them. That is, they are independent of any particular external environment, and of each other. Conceptual process *A* may be linked momentarily with conceptual process *B* at one time, with *C* or *D* at another. We are not talking now of the processes underlying the elaborate sentences of adults, but simply of what is needed to account for the performance of the child who can use a vocabulary of four words to form the propositions "Daddy go," "Mommy go," "Mommy come," and "Daddy come." This is beyond the level of any animal except man; but putting the problem in these terms allows us to relate it to other phenomena that do occur in the animal world.

Let us propose as hypothesis the more general proposition, that the intellectual development in phylogenesis that eventually makes speech possible is an increasing independence of the conceptual activity from the present sensory environment, and an increasing capacity for entertaining diverse conceptual processes at the same time. This is the capacity to respond to the present environment in one way and think of responding in another; for perceiving the present situation as it now is, and conceptually adding something else to it; for planning a series of actions of which only the first is immediately feasible. At the lowest level, it is the capacity for delayed response or a simple expectancy; at the highest level, for "holding" not only a series of words but also of sentences, whose meaning only becomes clear with later words or sentences (Hebb and Bindra, 1952).

There is clear evidence of phylogenetic development in such capacities. The chimpanzee moving a box to stand on, to get fruit that is otherwise out of reach, manifests a "first this, then that" planning which is beyond the dog, and Köhler's (1927) evidence leaves little doubt that the sequence of acts is, in fact, planned in advance. Entirely unmistakable evidence of such ability is found in the deceitful attack that almost any adult chimpanzee is capable of. We shall return to this behavior in a later section, but here we may cite the chimpanzees of the Yerkes colony who, seeing a visitor being escorted through the Laboratories, (a) slip over quietly to the faucet and get loaded, (b) return to the front of the cage nearest to where the procession will pass and sit quietly, with no sign of hostility, until (c) the visitor is within range, at which point he is suddenly drenched with a pint or so of water. This makes two points. It is evident, first, that a series of actions is planned in advance; and secondly, that the chimpanzee has the ability to act in one way while thinking about acting in the opposite way.

Closely related is the empathy described by Köhler (1927, p. 168 and Plate VII). One chimpanzee watches another precariously reaching for food. The watcher stays at a distance, does not try for the fruit himself, and still identifies with the other to the extent of making identical reaching movements. That such spontaneous imitation can occur in chimpanzees is confirmed by the socially significant paper of Hayes and Hayes (1951b), and Köhler's empathy should be collated with Mead's idea (Morris, 1946, p.

45) that communication involves "taking the role of the other" and "the ability to be the other *at the same time* that he is himself" (italics ours). Such a proposal becomes intelligible if two or more phase sequences (Hebb, 1949b), or trains of thought, can exist side by side in the brain of the higher mammal.

But this, of course, is not the place to elaborate a whole theory of communication. We can sum up the present discussion by saying first that the criterion of syntactic behavior clearly separates man from ape, in communication; but secondly, that the appearance of some other aspects of behavior in the chimpanzee shows that the ape may be closer to man in this respect than he seems.

ANIMAL COOPERATION

Much the same conclusions must be drawn concerning cooperation. Phyletically, cooperation parallels communication to such a degree that there is a *prima facie* case for concluding that both depend on the same intellectual characteristics. The cooperation of insects, as we have seen, is complex indeed but shows little sign of purpose. Short of man, the highest level that has been demonstrated experimentally is in the chimpanzee (Crawford, 1937, 1941) but it seems true of his cooperation, as of his communication, that the ape is not decisively superior to other mammals and definitely lacks a kind of teamwork that is common in man (Nissen, 1951). Crawford's experiments have been summarized by Nissen (1951, p. 444) as follows:

> In a laboratory experiment, Crawford first trained young chimpanzees individually to pull in a weighted box, by means of an attached rope, in order to obtain the food reward on the box. Then the weight of the box was increased (so that one animal could not move it); next, two ropes were attached, the box was baited with two portions of reward, and two animals were put in the cage. At first each chimpanzee pulled without any reference to his partner's actions; only by chance did the two pull at the same time and so move the box. It was necessary for the experimenter to give a signal, which the partners had individually associated with initiating the pull, to get them to pull together. Once the temporal coordination of effort had been established in this way, the extraneous signal could be omitted. One of the partners, *A*, usually watched the other one, *B*, and was ready to add her pull as soon as *B* started heaving on the rope. Sometimes it was arranged that *B* was satiated when the experiment began; *A* would then urge and direct *B* to pull. Eventually *B* would pull, although she had no desire for the food reward. In these circumstances *A* often reaped the entire material benefits of the teamwork.
>
> In a later experiment by Crawford one animal was trained to push against four colored plaques, arranged around the sides of a cage, in a certain order. The cage was then divided into two sections by a partition of bars; the colors to be pushed first and third were on one side, the colors to be pushed second and fourth were in the other part of the cage. The trained animal was then put on one side, and her partner, who knew the mechanics of pushing the plaques but not the correct order, was put on the other side of the partition. The trained animal *X* watched her partner's actions and pushed the color available to her after *Y* had pushed the color preceding it in the sequence; *X* sometimes solicited her partner, using begging gestures . . . that were spatially directed toward the particular color on her partner's side that was the next one in order.

We know of no reliable report of cooperation by other animals that quite comes up to this, with its evidence of purposiveness. Yet it must be pointed out that there was hardly more than the barest beginning of teamwork as we know it in man, and that even this was achieved only in an artificial experimental setting which needed the planning of the human experimenter. The experiment may be thought of as parallel to that of Hayes and Hayes (1951a) in communication, who managed by careful training to get a chimpanzee to produce recognizable English words, but still without achieving real language.

What may be described as a "one-sided cooperation" certainly occurs in the chimpanzee. An adult animal rescues an infant heading for trouble (Nissen, 1931, p. 87): *A* helps *B*, but *B* does not simultaneously help *A*. Teamwork at the human level is another matter. It involves a kind of prior understanding between cooperators that is impossible, perhaps, without a human capacity for communication. When *X* and *Y* cooperate fully, very complex mental processes are required. At each stage of action, *X* must anticipate not only his own next act and its effect, but also those of *Y*. Further, *X* may have to keep a number of possible eventualities in mind, and in each case have the expectancy of *Y*'s acting in an appropriate way. This is the relationship that Sears (1951) has defined as dyadic. It is clear that the expectancies involved become very complex when the situation is not under the complete control of the cooperating partners. Teamwork thus makes intellectual demands of the same order as those made by language. Psychologically, it may, in fact, be hard to distinguish the two.

Let us then define three levels of cooperation, corresponding very well on the whole to the three levels of communication already proposed: (a) reflexive, or nonpurposive; (b) purposive, but "one-sided"; and (c) "two-sided," or teamwork. The second of these is seen in higher vertebrates; the last seems to occur only in man, as far as one can tell from the available evidence.

The difficulty in reviewing the literature on cooperation in its various forms is that those reporting have usually not known or have not understood the significance of Tolman's (1932) clarification of purpose as a behavioral and nonteleological conception. They have tended therefore to see purpose in all behavior, since almost any act is bound to have *some* biological significance; or at the other extreme have tended to make all infrahuman social behavior reflexive and without any insight whatever. In either case the details of the behavior are often not reported in a way that allows one to judge whether its end effect is achieved with insight and intention or not. There is a great deal of anecdotal material to suggest the existence of altruism and cooperation in primate societies, for example, but the evidence in general has been so poor that Zuckerman (1932) had to conclude that no altruism exists, that the behavior is always "immediate responses to specific stimuli," without comprehension. We shall try to show that this is not so, and that one-sided purposive cooperation and even the beginnings of altruism can be found in lower animals, but it certainly seems that much which at first sight is a planned cooperation is no more so than a litter of kittens is planned parenthood on the part of a tomcat.

Lorenz (1934) has reported a good example of behavior that only looks altruistic. He found that jackdaws would always come to the aid of another daw that had fallen into the hands of a common enemy, like Lorenz himself. When he had a captive bird in his hand, the others would attack him vigorously. But they would also attack him, with equal enthusiasm, when carrying a wet black bathing suit. Assuming that the jackdaws' visual acuity is equal to the discrimination involved, as it must be, there is no reason to suppose that either attack was a purposeful attempt to help a fellow.

A further difficulty, in reviewing the literature on cooperation, is that some of the most interesting and significant observations are made in the field, and are naturalistic instead of experimental. The behavior cannot be repeated at will, varying the circumstances for analysis. The phenomenon of leadership is an outstanding example: leadership can be quite independent of dominance and has much greater social significance, as an integrating instead of disintegrating influence (Darling, 1937; Carpenter, 1942). But leadership cannot easily be studied in a cage, and dominance can. As a result, dominance has been greatly overrated as a social phenomenon (Schneirla, 1952), and despite such excellent reports as Darling's or Carpenter's we know little about the psychological conditions of leadership or the closely related role of the sentinel.

Dominance often determines leadership, but often not. Greenberg (1947), for example, found that dominant sunfish also tend to be the leaders in group maze-running. Fischel (1927), on the other hand, did not find such a relationship in hens. The leaders in foraging expeditions in an orchard were not always the dominant birds. Katz (1949) with Barbary sheep as subjects, Scott (1945) with domestic sheep, and Stewart and Scott (1947) with goats, report that dominance does not play a major role in group organization. Leadership, however, is of great importance and usually rests with a female. Such "matriarchal" societies are very different from the "patriarchal" societies of roe deer (Darling, 1937) or baboons (Carpenter, 1942), in which the most aggressive and dominant male occupies the central position of control, and Darling has stressed the value of matriarchy as making for true group cohesion. The leadership of the hind is more "selfless," whereas in a patriarchal society the dominance relation, between leader and followers, is directly at odds with the leadership relation. Thus the hind group among the red deer is much more closely knit than the stag group.

As we have said, not much is known about the way in which leadership is acquired, or the level of insight that is involved on the part of leader and follower. That is, we do not know whether the follower is less alert than the leader *because* the leader is already acting as sentinel, or whether the follower is simply a less alert animal. Is the leader more alert, as leader, than he would be as a solitary animal? Even in Darling's (1937) report of the greater selflessness of the leader, who did not go to food until satisfied that the coast was clear and who as a result got less food, it is not clear that concern for others was in question, or anything more than a special wariness.

FRIENDSHIP AND ALTRUISM

But we do not regard Darling's opinion as fantastic, or even as intrinsically improbable, although it must be hard to establish in the conditions of his study. There is definite evidence in other animals of a phylogenetic development of something we shall call altruism, defined as an intrinsically motivated concern for others. Without attempting to make the definition exact, we use "altruistic" as a convenient designation for certain behavior that is not reflexive or immediately rewarding biologically: not necessarily "generosity," nor acting against inclination, but due to a capacity — or weakness — for developing a nondestructive interest in others. The weakness is very marked in man, and it does not matter for our present purposes whether it is a "learned drive" (Miller, 1951) or not. If the concern for others must be learned, the comparative evidence still indicates that we are dealing with something fundamental to man's nature, for we can trace phyletically some signs of an increasing readiness for such learning and an increasing frequency of the behavior in question without special conditions of reinforcement.

The biologically primitive form of a concern for others is, first, in parental care, at a purposive level. Secondly, it is clearly seen when an attachment to a sexual partner extends beyond the actual period of copulation, a phenomenon which is said to occur in some birds and mammals.

A further step in development appears in selective friendships which do not depend directly on primary reinforcement in Hull's (1943) sense. Such friendships are occasionally seen between dogs, and we may cite also the pet dog's protective behavior toward younger members of the family, when they are not the ones who feed or pet him. The clearest evidence that this behavior can be independent of primary reinforcement is from the dog that makes a nuisance of himself by persistent "rescue" of children in swimming, despite repeated punishment.

With the chimpanzee one cannot fail to see a step closer to man. Most chimpanzees, most of the time, seem entirely selfish; but so often are people. Friendships are marked (Köhler, 1927), and as Nissen (1951) has shown, the chimpanzee may be upset and unhappy when the long-time friend is taken away, even though other animals are still present.

Nissen's description of the effect of begging is particularly interesting here. It seems to be "both unpleasant and a very compelling stimulus. . . . Often it appears that the [chimpanzee with food] cannot tolerate this; reluctantly or even with some show of irritation or anger, he hands over some of the food, or throws it violently at the beggar" (Nissen, 1951, p. 445). From one point of view this is not generosity at all, but from another the very reluctance to give makes the gift all the more significant, by showing the strength of the compulsion behind it.

Here we can also give an example of what seems a purely altruistic concern for another chimpanzee, in an attempt to get the other out of danger. In the "bold and timid man" experiment described in the following section, Lia was afraid of the experimenter (outside the cage) playing the part of the bold man, and usually preferred to stay away from him. But Mimi, with whom she was caged, had other ideas. She stayed close to the cagewire and attacked repeatedly. The role of the bold man demanded that aggression be met with aggression, and the following note was made on the data sheet.

> Mimi tried to get my hand into reach, obviously to repeat her earlier attacks. I caught one of her fingers with my gloved hand and bent it back slightly, enough to hurt, whereupon she screamed with rage, beat on the wire, and then returned to putting her hands out through the wire, trying to catch hold of me. Lia, in the meantime, was also angry at first, when I hurt Mimi, and attacked me until I struck at her hand with my fist. Thereafter she was alternately aggressive in my direction (but not getting within reach) and protective of Mimi, five or six times *pulling Mimi's hands back from the wire,* first one, then the other, to get her out of danger (Feb. 14, 1944).

A week later when the next bold-man test was made, Lia showed no anger or aggression but, when the foolhardy Mimi persisted in trying to get the bold man into position for an attack, strenuously tried to pull her away for the whole observation period of one or two minutes.

Finally, we may conclude with a remarkable observation of friendship between porpoises which is reported by McBride (1940, p. 26) — remarkable especially because the female referred to was sexually receptive, and before and after the incident described the two males were each bent only on preventing the other's access to her:

> The smaller of the two large males was removed from the tank and transported by means of our tank truck to Tampa. It was exhibited there under the auspices of the State Conservation Department in a concrete tank for three weeks. When the animal was released into the tank, the greatest amount of excitement on the part of the larger male was exhibited. No doubt could exist that the two recognized each other.

> and for several hours they swam side by side rushing frenziedly through the water, and on several occasions they leaped completely out of water. For several days, the two males were inseparable and neither paid any attention to the female.

Socially, the porpoise is an interesting animal indeed, and it is most unfortunate that naturalistic observation outside captivity is impossible. On the one hand, dominance is clearly marked, and vicious attacks with no apparent reason have been repeatedly seen. On the other, very elaborate and apparently unmalicious teasing and social play are prominent in the animal's behavior, as well as clearly marked cooperation. At sea in their own habitat the female presumably gives birth in isolation; at any rate, the female giving birth in captivity in the Marine Studios was very exciting to the males, one of which became aggressive. The other females gathered round the one in labor and helped ward off the attacking male. When the newborn infant began its first gradual ascent to the surface to breathe, another female accompanied the mother in swimming just below the infant in readiness to support it if it had failed to make the grade (McBride and Hebb, 1948). From other behavior in these animals it is reasonably certain that the behavior described was purposive, to assist the infant if necessary to get to the surface (as it must, of course, to breathe), for another mother whose infant died during the night was found in the morning steadily supporting the dead body at the surface. If such a level of behavior at first seems surprising, it may be remembered that the porpoise is a mammal with a brain fifty percent bigger, and with more cortex and a more convoluted cortex, than man's. The brain lacks a frontal lobe, but in other respects it has striking similarity to that of the higher primate (Grünthal, 1948). It is not anomalous, accordingly, to class the porpoise *(Tursiops truncatus)* as a higher animal (McBride and Hebb. 1948).

The evidence indicates, therefore, that a disinterested concern for others can be found in mammalian development. Although it reaches its greatest potential only in man, it is not foreign to any higher mammal, and is not something that is imposed only by reward and punishment on the growing human child. It can, of course, be stultified or fostered during growth, but by his intellectual and emotional characteristics man has a greater *aptitude* for altruistic attitudes than any other animal, and it is of great importance to gain a more precise knowledge of the conditions of its development. It is not likely that such knowledge will be obtained by human study alone (who would attempt such an experiment?), but here again there is hope that real advances will be made in pooling the methods of animal experiment and naturalistic human observation.

The Analysis of Social Attitudes

We have argued that there is a definite gap between man and his nearest relatives in their effective levels of cooperation and communication. No chimpanzee has language, or takes part in teamwork as we know it in man. But it was suggested also that the chimpanzee may not really be as far from man as he seems in these matters, and we turn now, in this and the following section, to behavior in which he seems human indeed. The evidence with which we deal here concerns, in general, the chimpanzee's social motivations; and the social situations considered will be those particularly which comprise two individuals, the dyadic unit of Sears (1951): chimpanzee-chimpanzee or chimpanzee-caretaker.

We also come now to things that have more obvious meaning for human social psychology. In this chapter, one of our objects, of course, is to persuade the reader that there is such meaning in animal work. Since it is most evident and direct with chimpanzees, it is here that most of our emphasis must lie; and this in turn implies giving most of our attention to observations made at the Yerkes Laboratories of Primate Biology, the chimpanzee breeding and research station in Orange Park, Florida. The work at the Yerkes colony is of unique importance for psychology, especially in the general area of personality or temperament. Not only does it constitute the only really extensive body of work on the great apes (Köhler's was a single-handed attack, now nearly forty years ago), it also constitutes the only thorough longitudinal study, from birth onward, of a higher mammal, the same animals being studied by a number of persons from a variety of points of view (Yerkes, 1943).

It has been noted elsewhere (Hebb, 1946a) that exposure to a group of adult chimpanzees gives one the overwhelming conviction that one is dealing with an essentially human set of attitudes and motivations. The feeling is not less strong in the psychologically trained than in the untrained (and its strength can be embarrassing to the purist who thinks that naming an animal's attitude or emotion is "subjective" and commits the deadly sin of anthropomorphism).

It is probably a common experience to all who have worked at the Yerkes Laboratories to feel that the bare bones of human personality, the raw essentials, are being laid open before his eyes. At the same time, it is hard to convey this to others, and to support it with behavioral evidence.

Merely describing what the chimpanzee did that gave one this feeling is not enough. One would also have to take the space to put the act into its setting: a life history of the animal, plus a detailed description of his usual habits. As the reader will see, interpretation of the act depends on such a background. The interpreter usually does not realize how he arrives at the interpretation, and so is embarrassed to find himself holding what seem to be unscientific convictions; but it is possible to show that the answers, though highly inferential, are based on objective evidence of the same kind that is used in identifying human attitudes and emotions (Hebb, 1946a). Even though such studies have not gone far, therefore, they have significance for understanding human social interaction.

Some of the chimpanzee behavior with which we are concerned takes the form of rather dramatic episodes that are infrequent and cannot be repeated at will for study. These verge on the anecdotal, perhaps, but cannot be left entirely without mention when one wishes to show at what social level the chimpanzee operates, at least on occasion. For example, May more than once convinced competent observers that she was skillfully and intentionally transforming the hostile pursuit of a male, bigger than she, into a friendly game of chase. Another example is Lia's attempt to keep Mimi from a foolhardy attack on the "bold man," cited in the preceding section. Perhaps the most striking of these isolated observations is Vera's deliberately picking a fight with Jack in the next cage in order to distract Shorty's attention from the food being brought for them both — Shorty and she being caged together, Shorty dominant and sometimes taking her food. The interpretation in this last case rests partly on the fact that Vera's hair was not erect (as an excited chimpanzee's would be). Also, she left the fight as soon as Shorty entered it on her side and came without evident excitement to get and eat her food. Finally, she screamed loudly again, without leaving her food, when the fight between Jack and Shorty began to peter out without her help. The scream renewed it for the better part of a minute.

Such reports, being reports of interpretation as well as of fact, do not, of course, make very satisfactory evidence. The objection may apply less strongly to some of the repeatedly recorded aspects of the chimpanzee's sexual behavior. Examples are Pan's preferential interest in a young female with whom he had never copulated, so that he masturbated while she was led by out of his reach and disregarded the fully receptive female within his own cage; and Fifi's repeated refusal to copulate with Bokar over a three-month period (three periods of oestrus) though she had mated satisfactorily with him on earlier occasions and though the two were on friendly terms in other respects. Placed with Frank, Fifi then mated at once, quite normally. Such a selectivity has been observed repeatedly, with a number of animals. A refusal to copulate may or may not show social intelligence, but it does show how far the chimpanzee's sex motivation is from the more reflexive (hormone- and sense-dominated) behavior of dog or cat. It shows, in short, the dominant role that attitude may have in the chimpanzee's social behavior, and as such a closeness to man.

Yerkes and Elder (1936) have given a detailed description of the chimpanzee's mating behavior, and the reader will also find accounts of a fascinating interplay between sex and dominance in the chimpanzee in papers by Yerkes (1939, 1940, 1941), Crawford (1940), and Clark and Birch (1945; Birch and Clark, 1946, 1950). In general, there is a clear relationship between phylogenetic level and a developing complexity in sexual behavior which Beach (1947, 1948) has worked out in detail, and which is applied by Ford and Beach (1951) to a clarification of human sex behavior. Their last chapter in particular, "Human Sexual Behavior in Perspective," is an excellent example of the contribution that a comparative study can make to an understanding of man.

RECOGNITION OF SOCIAL ATTITUDES

The conditions in which social attitudes are identified in the chimpanzee allow one to undertake a kind of analysis that is rather rare in comparative psychology. That is, while pretending to study the chimpanzee, one can really be analyzing the reactions of the caretaking staff (including one's own reactions) and ask, "Just what cues in the chimpanzee's behavior determine the human judgment that he is interested, or friendly, or malicious?" With the animal safely behind cagewire, unable to bring suit for damage to his self-esteem, one can experiment with emotion better than if the subject were man, but one can also study better the human recognition of emotions, with fewer confusing variables and more complete access to the data affecting the observer's judgment.

If the judgments were made at random, the analysis would have little significance. But it will be found that the experienced man refuses more often than not to say what the motivation is behind a given act. The behavior is often very puzzling to the caretaker, but when he names an attitude or motivation with confidence it is generally a dependable guide to the animal's future behavior. The naming is not made from superficial similarities with man, which are often very misleading, and the worker may want to know an animal for years before venturing an opinion as to what lies behind some particular act (Hebb, 1946a).

This provides an exceptional opportunity for study of the complex social stimulus. What are the actual cues that determine a social response? What parts of the response determine the further responses of others? When a child acquires some attitude from his parents we neither know what stimuli entered into the learning, nor how the learning occurs. We know little about the effective cues by which an emotion or attitude spreads infectiously through a crowd. As Allport (1940) and Jones (1944) have pointed out, it is urgent for psychology to find out what happens in the common-sense intuitive social judgment. Until that is done, there remains a fundamental defect in the data of social psychology. One could almost wish that the various rating and ranking methods had not been devised, for their success in quantifying social judgments has allowed us to forget that we do not know what is being quantified and has led us to neglect the essential task of analyzing these implicit predictions of human behavior so that we can learn to improve on them.

The first key fact is that recognition of a "mental state" in the chimpanzee usually does not depend on the present behavior alone. It is a perception of the relation of the present behavior to what the animal does in other circumstances. "Interest" describes a heightened level of responsiveness, not a particular level: the behavior that is a sign of interest in one animal is apathy in another. "Anger" implies a sharp deviation of behavior from an accustomed base line.

Accordingly, such identifications usually require a thorough, long-time knowledge of the individual, in addition to background knowledge of the species. Experienced caretakers refused, for example, to make any statement about the attitudes of two chimpanzees, Shorty and Vera, with whom they had worked for some months, on the ground that they did not know them well enough yet. The animal's history seems crucial always. A single observation of aggressive behavior on the part of Frank, in rather unusual circumstances, caused a serious revision of the generally held view that he was friendly to man, and these doubts persisted despite Frank's exemplary behavior afterward. At one time Dita had not attacked a member of the staff of the Laboratories for several years at least, but a knowledge of some of her earlier behavior still made her distrusted.

On the other hand, animals like Bimba and Fifi were frequently aggressive, but their aggressions were put down to brief bouts of anger, again to a great extent on the basis of their histories. The judgment that an animal is basically friendly, therefore, is an especially complex inference; this view is supported by the fact that there may be some disagreement among the staff as to the friendliness of a particular animal, and that the judgment is often made very tentatively. Although it is definitely based on behavior, and thus "objective," it is hard to find the crucial factors that determine it.

But other judgments are more amenable to analysis. Let us present next a method of recording used in a study of individual differences, or of temperament, in the chimpanzee (Hebb, 1949a).

A METHOD OF RECORDING SOCIAL ACTIONS

Systematic samples of the behavior of 30 adult chimpanzees were taken during the course of the ordinary routine of the Yerkes Laboratories, and records taken also of behavior in the "bold and timid man" experiment, to be described shortly.

Describing the behavior was the essential problem. It would have begged the question completely to record that the chimpanzee was friendly or frightened or bluffing: such judgments were the ones to be analyzed. It was also very difficult to avoid contamination of the data by such judgments made implicitly. Eventually some 36 specific acts ("first-order categories" of behavior) were defined in which interpretation was at a minimum, such that all observers could agree in their use. All records were made in terms of these acts, and thus seemed fully "objective."

At this level, however, the raw data were an indigestible mass in which no order or meaning could be found. The next step was to group together acts that had roughly the same behavioral significance. These "second-order categories" were "friendly behavior," "aggression," "quasi-aggression," and "avoidance" (together with a class of "no response"). Interpretation is definitely involved in making the classification, but it remains under control, and the data are recorded in first-order categories only. "Friendly behavior" does *not* mean "friendliness" but is defined as the occurrence of one of the following actions: putting the head, shoulder, back, belly, or arm against the cagewire at the observer's approach, or (if it was there already) keeping it there while the observer scratches or rubs the skin; thrusting fingers or mobile lips through the two-inch wire mesh (a common invitation to contact); grooming the observer's skin; and making a sputtering noise with the lips (Bronx cheer) while approaching the observer. "Aggression" and "avoidance" are self-explanatory; "quasi-aggression" is a combined category of intimidating and nuisance behavior: walking erect with the threatening swagger of the chimpanzee about to attack (or trying to bluff), beating the ground or some noisemaking surface, throwing sand or feces and spitting.

Order could now be found in the data, but it was still clear that something was missing: two animals that were known to be quite different might show quite similar totals in these higher categories. With the next step, however, light began to break through. The "third-order category" of behavior was defined as (a) the occurrence of behavior from two second-order categories in definite sequence (e.g., avoidance followed by aggression is one category, aggression followed by avoidance another) or (b) the occurrence of an act in one second-order category in the absence of any act from another specified second-order category (e.g., aggression not preceded by quasi-aggression). When the data were arranged in this way new meaning began to appear.

The change in analysis made by the third-order category is most clearly shown in the data for males versus females, in their relationships to the staff. First consider the second-order categories of friendly behavior, quasi-aggression, and aggression, as in Table I. These data say that male and female show friendly behavior about equally often, but that the male is significantly more aggressive. The implied conclusion, Look out for the male, is markedly out of line with the caretakers' experience; nineteen out of twenty cuts or scratches they receive come from the female. Now consider the third-order categories of *aggression preceded by quasi-aggression* (the openly prepared attack), and *aggression preceded by friendly behavior* (deceitful attack), as in Table II.

Since Table II may seem to have an antifeminist bias, we hasten to add several things. The male does not appear from the data to be less often desirous of doing injury than the female, but may be less astute about it. The deceptive attack on the staff by males does occur, although rarely enough that it was not found in the behavior sample; and 7 deceptive attacks were made by males in the timid-man experiment (mean, 0.9; for females, 1.4). Some females, also, tend to show a male directness of attack, and individual differences in this respect seem to be greater than among males.

The method of analysis also clarifies certain judgments about individual animals which would puzzle an outside observer, or perhaps confirm his suspicion that the one making the judgment is going far beyond the facts, when a caretaker says that exactly the same behavior in one animal is the expression of anger and in another attention-getting. The judgments are complex, but one can now show that they are based on observable behavior.

Take the simple case in which the same act, spitting, is called evidence of *desire for attention* in one animal, *malice* in another. At the lowest level, we are dealing with two sequences of behavior, spitting followed by friendly behavior and spitting followed by aggression. From past observation of each animal the observer can

TABLE I

INCIDENCE OF THREE FORMS OF BEHAVIOR BY MALES AND FEMALES: SECOND-ORDER CATEGORIES

		8 males	*22 females*
Friendly behavior,	total . . .	205	531
	mean . .	25.6	24.1
Quasi-aggression,	total . . .	73	25
	mean . .	9.1	1.1
Aggression,	total . . .	41	38
	mean . .	5.1	1.7

TABLE II

INCIDENCE OF TWO FORMS OF BEHAVIOR BY MALES AND FEMALES: THIRD-ORDER CATEGORIES

	8 males	*22 females*
Quasi-aggression, then aggression	37	0
Friendly behavior, then aggression	0	15

predict the second part of the sequence and as soon as the spitting is seen, without further observation, can say, "That is (or is not) malicious." The diagnosis is a prediction, and the prediction does not arise from some esoteric penetration into the consciousness of the chimpanzee but from a knowledge (a) of what chimpanzees in general are capable of and (b) of what this chimpanzee in particular has done in the past. Again, *bluff* is the sequence of intimidating behavior followed by retreat if the opponent does not retreat but moves forward. As soon as he sees the intimidating behavior in certain circumstances the observer may call it "bluff" without having to see what follows. But he can do this only from prior observation, so his diagnosis once more is a prediction.

The reader may think all this too simple to need to be spelled out. In practice it is not. One's *present* perception of behavior is deeply colored by past experience. Experimentally, one may waste months in a futile effort to identify subtle distinguishing marks that are not there.

For example, it usually seems to the observer that there is something in the manner of the bluffing chimpanzee that is different from the manner of one that is really going to attack. Perhaps there is, but the experienced handler of chimpanzees still declines to make the judgment with confidence until he has seen a good deal of the animal, which indicates that he does not really identify the bluff by present behavior alone. One of us (Hebb) had worked on this whole problem of analysis for three years before he discovered that he was still recording intimidating behavior in two ways, according to what he thought it meant: he recorded "bluff" when aggression did not follow immediately, but "aggression" when it did. In the latter case the prior intimidating behavior did not seem something separate—it was a natural part of the whole picture of *one* aggressive act. The record of behavior became clear only when the two patterns were put down as (1) quasi-aggression followed by aggression, and (2) quasi-aggression not followed by aggression. Bluffing is a third-order category, that is, not first order. Bluffing, as such, should not be entered in the record of behavior.

The same keen observer took two years to discover that his observations of Bimba's attacks were colored by his belief that the behavior was caused by anger, not cunning malevolence, so that he failed to see that the attacks were sometimes preceded by friendly behavior to get the victim within reach. Only when a satisfactory method of recording was used did the truth appear.

THE "BOLD AND TIMID MAN" EXPERIMENT

Any sensible person working with adult chimpanzees is constantly on guard, even with animals that in general are well disposed. But this very caution may be an incitement to attack, and it seemed possible that further light on individual differences might be had if one could test the animals' reactions to a person with no fear, on the one hand, and to an exceptionally fearful one on the other.

Accordingly, the two roles were set up, played by the same person in two disguises. In the one, with hands protected by heavy gloves, the experimenter strode confidently up to the outside of the cage, making rough but friendly contact when the animal stayed within reach without aggression; but any attempt to scratch or bite was answered by striking back. The timid man acted the part of an excessively nervous person, very hesitant in approach, starting back at any sudden move of the animal, and making no near approach to the animal as long as he was aggressive in manner.

TABLE III

PERCENT OF TRIALS (2700 WITH STAFF, 450 WITH TIMID, 450 WITH BOLD MAN) SHOWING CERTAIN RESPONSES: Fr, FRIENDLY BEHAVIOR; Ag, AGGRESSION; QAg, QUASI-AGGRESSION; Av, AVOIDANCE (DATA FROM HEBB, 1949A)

	Staff	*Timid*	*Bold*
Apparently friendly (Fr alone)	26	29	23
Fear (Av alone)	0.2	7.1	37
Aggression (Ag in any combination: 2nd order category)	5.0	13	13
Annoying or teasing (QAg without Av or Ag)	3.7	24	3.8
Bluffing (QAg-Av)	0.0	2.2	5.6
Deceit (Fr-Ag or Fr-QAg)	0.9	17	2.7
Deceitful attack (Fr-Ag)	0.6	8.7	2.2
Deceitfully startling or annoying (Fr-QAg)	0.3	8.4	0.4
Afraid but friendly (Av-Fr, Fr-Av)	0.0	2.7	2.7
Afraid, antagonistic (Av plus Ag or QAg)	0.0	2.7	10

Some of the results are given in Table III. The outstanding one was fear of the bold man, although he did nothing to start trouble but only answered in kind. Avoidance alone occurred on 37 percent of trials with the bold man, as in the table, but if the trials in which avoidance was combined with some other behavior were included the total would be 49 percent. Some of the fear of the bold man presumably was due to the strange costume worn, since the rather similar appearance of the timid man also aroused some fear (bold and timid man each wore false faces) ; but the two costumes were similar, so it may be concluded that the difference between 37 percent avoidance of the bold man and 7 percent of the timid man was due to the difference in behavior. It will be recalled that both roles were played by the same person.

A more significant result for our present purposes is the high frequency of aggression and quasi-aggression to the timid man, and especially the frequency of behavior which, to state the matter as clearly as we can, enticed an innocent close just to scare the hell out of him. The role, of course, demanded that the timid man fall into this trap as often as it was laid, and that he should show a maximum of startle every time. Many animals found the temptation irresistible; only 9 out of 30 failed to show such behavior on one or more trials.

The individual differences revealed by the test method are at least as interesting as the group results, although they cannot be gone into thoroughly here. Pan, often aggressive to the staff but considered a gentleman because he always gave clear warning, began by attacking the bold man; being outbluffed (thanks to a strong cagewire) , he appeared to lose all interest and moved off to a far corner of the cage whenever the bold man appeared. Two much more vicious animals, Pati and Beta, were equally impressed by their failure to cow the bold man, but in a different way. Like Pan, they first attacked; when the bold man did not retreat but struck back, they became markedly friendly. To the observer this transition was little short of astonishing. Instead of their usual lack of interest in man, alternating with vicious attack, they now vigorously solicited friendly contact. The behavior was not fawning by any means (both tried occasional further attacks, for example) , but they appeared to be strongly impressed by the bold man, and on the whole very favorably.

Another pattern of response was shown by Mimi. She was markedly friendly to the timid man, and was tempted into only one attack on him, midway through the experiment; but she attacked the bold man with persistence and determination. Still other responses were shown by Jack and Frank, who were highly interested but unaggressive, and unafraid, with both bold and timid man.

THE CONDITIONS OF FEAR AND HOSTILITY

The results of the experiment just described serve very well to introduce our next main point, which is that the social determinants of fear and hostility are very complex. If the reader does not find this conclusion new and startling we can say it another way. The problem of social motivation is apt to be taken too lightly, with too much of the reversed-panacea approach—the idea that all our social troubles stem from economic pressure, *or* Oedipus difficulties, *or* frustration, or some other single source. Fix whichever one it is that makes the trouble, and all will be well. A main object here is to show that the infrahuman animal has a more complicated set of motivations than is allowed to man by these one-shot prescriptions.

Some of our conclusions in this chapter will be quite obvious ones: frustration is not a sufficient statement of the causes of aggression; the unknown, as such, may cause fear; man, who so often avoids pain and effort, also goes to some trouble to expose himself to threat of pain and to situations demanding effort. These things are common knowledge. But they are not common knowledge in psychology, and we must remind the reader that everyone knows a lot of things that are not true. One necessary task for scientific psychology is to sort out, among the obvious facts of human nature, the true from the untrue, the useful ideas from the useless. The reason for preferring simpler theories of human action is not that anyone thinks that the facts are really simple, but that it is too easy to postulate explanatory entities whenever the going gets hard. Which are the fertile postulates? Which are really needed to account for the facts? It is especially necessary to be careful when a discrepancy between fact and theory is found only at the human level, in the adult subject, because of the multifarious and uncontrolled experiences that occur during the long period of human growth, together with the special effects of verbal learning. So it will not be new if we conclude from the animal data that man is hostile to strangers; this is "known" already, but the need for taking theoretical account of it is clearer if its phylogenetic development can be traced, and when related behavior can be demonstrated in lower animals, in more controlled conditions and without the contaminating influence of language.

CHIMPANZEE FEARS AND AGGRESSIONS

Fear of innocuous objects is very common in the chimpanzee, even though these objects are not associated with some more primitive cause of avoidance. The diary records of the Yerkes Laboratories contain a great many references to unexplained fears of a particular person, some piece of apparatus or part of it, rope of a particular size, color, and texture but not of other ropes, a biscuit in which a worm is found, a vegetable of a peculiar shape, and so on. There are great individual differences, of course, but as a group the adult chimpanzees are markedly subject to such fears. Köhler's (1927) observation of panic at sight of small animal models such as a toy donkey is well known. A chance observation, made with a clay model of a chimpanzee head (Hebb, 1946b), extended the range of objects known to have such effects. The clay model, reasonably lifelike in proportions but about half life size, produced extreme panic in nearly half the thirty-two adults to whom it was shown, and very marked signs of fear in most of the rest. An actual head preserved in formalin had the same effects, and so had a very lifelike colored model of a human head from a display dummy.

An attempt to deal with such facts theoretically (Hebb, 1949b) is not too satisfactory in certain respects, but the line of reasoning suggested a relationship to human emotional disturbance at sight of a mutilated body and fear of the dead, and so the following observations were also made. These are cited in full because of the socially significant aggressions that were elicited along with the fear. A dead chimpanzee was not available for experiment, but anesthetized ones were:

> *Aug. 15, 1943.* Mars, a young animal who had been anesthetized by nembutal, in order to make physical measurements, was carried out and shown to four of the adults, Jack, Dick, Don, and Dita. The first three of these were markedly excited (Dita less so) and Don attempted to attack. *Aug. 17.* Mars, anesthetized again, was laid on the ground in the infants' enclosure, with his fellow youngsters. One of them, Art, was persistently aggressive, biting and hitting the inert animal, who had to be rescued. Mars was then carried to the cages of some of the adults. Nana and Helene showed interest but no emotional excitement; Don was excited, with hair erect, and somewhat aggressive in manner; Dina and Bula manifested a nonspecific excitation, with hair erect; Wendy's hair was erect, she did not seem as excited, but she came close and remained; and Josie (as it happened Mars' mother, from whom he had been separated early) manifested an apparent mixture of aggression and fear — she was very excited, ran inside when Mars was first brought near, then attacked him through the cagewire. Her hair was erect throughout.
>
> Since the sight of an anesthetized infant being carried by one of the staff was common, the observations above were later repeated with an anesthetized adult. *Feb. 11, 1945.* Don, under nembutal, was wheeled on a handcart up to the cages of nine other adults. Ami, Nira, and Vera showed fear, and Dina and Bokar did also but then followed this by a show of aggression at a distance; Kambi showed generalized excitation and screaming only; Frank, with hair erect, spat at the anesthetized Don; Pan first avoided, then attacked through the cagewire; and Lia, with general excitation but not avoidance, also attacked. (The youngsters in the infants' enclosure were afraid, one very much so, and all showed signs of marked excitation.)

Now let us look at the chimpanzee's aggressions as they occur in more ordinary circumstances, in the form of teasing, jealousy, anger, or chronic malice.

The motivation for teasing is often strong. It seems certain that the chimpanzee gets satisfaction simply from the discomforture of the human being who is startled and frightened, or angered, by a sudden screaming attack on the cagewire nearest him or on being squirted with water. Certainly the animal goes to considerable trouble to achieve such results, in spite of being often punished for it and getting no other reward than seeing the effects of the act. Other chimpanzees are teased as well as human beings, as Köhler (1927, pp. 83ff) noted. Deliberate planning is involved, with an evident expectancy of the end effects of the behavior. "Teasing" is too mild a word for some attacks that seemed to be made only for the sake of others' discomforture (as when Alpha would occasionally yank a handful of hair out of a thoroughly dominated cagemate, with no provocation), but it is important to note that these minor aggressions are not always indications of a desire to do more serious damage if only the victim were within reach. Quite often he *is* within reach, or becomes so later without receiving more damage. In the timid-man observations, some of the animals seemed genuinely malicious, but others, after once getting a rise out of the experimenter, became gentle and made no attempt to injure him when he came within reach.

Look next at the causes of anger, and their relation to frustration. The usual stimulus to anger is the behavior of a person or another chimpanzee acting in a way that is not desired, or failing to act in a desired way. "Desire" is one of the terms proscribed in Watson's Revolution along with images and ideas, not because of its vagueness or trouble in defining it (after all, "learning," "intelligence," and "emotion" were retained) but because such terms were thought to be entirely subjective and anthropomorphic in their reference. We can now see that they are much better treated as referring to intervening variables, postulated processes, either with man or animal. As such they are more or less suitable for use in an objective psychology. They are still vague, but the fault is not in the way the terms are defined but in our theoretical conceptions, which will not be made any better just by using new terms. If one is really serious about stating the conditions in which anger occurs, a reference to desire or some equivalent is necessary.

To speak of "frustration" instead does not avoid the necessity, for it is clear that it is a *desire* or an *expectancy* that is frustrated. By some writers, of course, the term "frustration" has been used to emphasize the complexity and importance of internal representative processes in a Freudian or near-Freudian way (see, e.g., Rosenzweig, 1944). With this we have no quarrel. But some writers have seemed to feel that the reference to frustration is a means of simplifying the problem, and that it avoids the need of postulating conceptual processes. Others have pointed out that this will not work for man (e.g., Sargent, 1948; McKellar, 1950). The evidence seems even clearer for the chimpanzee.

When a female chimpanzee in heat is introduced into the next cage, the male seeing her gets up and comes as close as he can, until he is stopped by intervening cagewire. At this time he shows no anger, although a goal-directed response, strongly motivated, has been interrupted. But anger becomes highly predictable if the male is first led to expect, by the caretaker's actions, that he will be admitted to the female's cage, and then is not. It is the expectancy that gives the physical barrier its effect. In other situations, one can only say that the chimpanzee is angry because he does not *like* what another is doing. No goal-directed response is interrupted, and a physical barrier has nothing to do with the cause of anger (although it may prevent the effects of the anger from being visited on its object).

Thus Pan had a temper tantrum at the sight of two experimenters taking a vaginal smear from a female in heat in another cage. Mimi was driven to paroxysms of rage by the two inmates of the next cage, who would wait till her back was turned and then spring screaming at the intervening cagewire, repeatedly startling her. Bokar was almost as much infuriated by Dick's spitting water at him, even when he was not hit. Köhler (1927, p. 293) has noted that noisemaking, and social disturbance as such, may be a cause of anger for a chimpanzee just as for man. The diaries of the Yerkes Laboratories likewise record a number of instances in which a noisy temper tantrum by one animal has led another to attack him. We have already commented (in discussing "altruism") on the anger of the chimpanzee at another's begging, even as he yields and hands over food. Defining the frustration in such cases of anger will present difficulty, unless it is frankly incorporated in a theory of thought so that frustration can be treated as some sort of conflict with the thought process. With anger, as with fear, the phenomena of emotion point to the intimate relationship of motivation to intellectual processes, and the impossibility of treating these things in separate compartments of theory.

THE PRACTICAL PROBLEMS OF SOCIAL MOTIVATION

It is reasonable, even desirable, to oversimplify theory in its preliminary stages, as long as one is concerned only with theoretical problems. The idea is that later stages of theory will do justice to the facts if early ones cannot. But it is quite another thing to forget that theory is still in these early stages when one is advising on a practical problem. What is virtue in one endeavor is vice in the other, and the more difficult the problem the truer this becomes. No other scientist deals with issues of such practical urgency as the social psychologist, and even though he has recognized the complexity of his problem the social psychologist may still find it worth while to consider some of the practical implications of the emotional behavior of higher animals, and perhaps also to be reminded, by the animal data, of the practical dangers of tacitly accepting too simple a view of motivation.

This last remark, be it noted, is not a covert attack on the theory of biological rewards (plus conditioning). Although we have probably betrayed our preference for another sort of approach, such as that of Harlow (1951) or Maslow (1953), we must point out that biological-reward theory is being greatly elaborated by such recent treatments of secondary reward as those of Meehl and MacCorquodale (1948, 1951), Mowrer (1950, 1952), and Miller (1951). These discussions imply no simple approach to the problem of social motivation.

In the year 1950 the main explanations of social hostility, as outlined by Katz (1951) in a review of the literature, make the hostility learned, *or* in some way a product of frustration. Although the difficulties in the way of one or the other of such simplifications have been recognized (Katz, e.g., p. 155, concludes in favor of multiple causation) the comparative psychologist may point out that the literature reviewed by Katz predicates a view of motivation that on our present knowledge is doubtfully adequate for the laboratory rat and quite inadequate for the dog or chimpanzee, let alone man.

From the practical point of view in this matter, the animal evidence offers good news as well as bad. It complicates the issues theoretically, but some of the complications mean that there may be hope for man as a social animal. On the plus side there is, as we have seen, evidence of a fundamental aptitude in higher animals for friendship and helping others. We do not know just what conditions of rearing foster this altruism, but it should be possible to find out experimentally.

A further plus item is that frustration does not always increase hostility, and may even decrease it without evidence of undesirable side effects ("displacement"). Punishing an adult chimpanzee for a misdemeanor like spitting or scratching the caretaker does not usually produce signs of resentment. On one occasion at least, punishment of a young animal for repeated biting led to a strong friendship (diary of Gamma, Yerkes Laboratories, 1935). The reader may be reminded also of the sudden friendliness of Beta and Pati to the bold man, when he promptly answered attack with attack. In animal, as in man, the effect of aggression in some circumstances is *respect* and *liking*. Nissen (personal communication) points out that forcing a chimpanzee to do something, or forcibly depriving him of something, is as often followed by an increase of friendliness as by a decrease. The friendliness cannot be an effect of the frustration alone; it must depend also on the fact that the frustrater is a rewarder at other times. But with a background of beneficence, the important thing is that deprivation need not decrease goodwill. Social living and education inevitably mean an unending series of frustrations but these, it seems, should not be thought of as always making for hostility.

But the bad news must be reckoned with, too. On the minus side is the evidence that timidity, as such, may incite aggression. Fortunately, this is not always so, and the consistently friendly behavior of some chimpanzees to the timid man justifies some hope concerning primate behavior in this respect. Also in the debit column are the mammal's fear and hostility with strangers, even when no injury has ever been received from a stranger (Hebb and Riesen, 1943). This has an evident meaning for social problems, and so has the combined fear and hostility toward a fellow chimpanzee under anesthesia (if, as it seems, this means that the cause of aggression was the failure of the anesthetized animal to behave in the normal way).

There is also another kind of evidence, which has both good and bad implications. It shows, in brief, that mammals seek excitement, and this search has desirable possibilities in society as well as extremely undesirable ones. The animal that is so vulnerable to emotional disturbance nevertheless seeks the situations that produce emotional disturbance in mild degree. There may be difficulty about dealing with this theoretically (Hebb, 1949b, pp. 232ff), but it appears empirically that for any one animal there is an optimal level of fear or frustration. First, the evidence concerning fear: strange surroundings tend to produce emotional disturbance in the rat, but Montgomery (1951, 1952) and Thompson (1954) have shown experimentally that the rat which has the choice of familiar and unfamiliar territory will tend to move toward the unfamiliar — the well-known exploratory drive. Whiting and Mowrer (1943) and Berlyne (1950) have suggested a connection between fear and investigative tendencies, and the dog that is frightened by a strange object is nevertheless apt to return to look at it again, balanced between closer approach and flight (Melzack, manuscript in preparation). The same thing can be observed in chimpanzees, and Woodworth (1921) and Valentine (1930) have described the behavior of young children who ask to be shown again — at a safe distance — the object that has frightened them. Secondly, concerning frustration: Harlow, Harlow, and Meyer (1950) have demonstrated the monkey's willingness to expose itself repeatedly to the frustrations inherent in problem solving, without extrinsic reward. Mahut, working in the McGill laboratory, has shown that the rat which is offered two routes to food, one short and direct, the other via a maze problem, will choose the problem on twenty to forty percent of the runs. *Some* of the time even a rat prefers to work for his living.

Such phenomena are, of course, well known in man: in the liking for dangerous sports or roller coasters, where fear is deliberately courted, and in the addiction to bridge or golf or solitaire, vices whose very existence depends on the difficulty of the problems presented and an optimal level of frustration. Once more, when we find such attitudes toward fear and frustration in animals, we have a better basis for supposing that we are dealing with something fundamental if a man prefers skis to the less dangerous snowshoes, or when we observe an unashamed love of work (problem solving and frustration included) in the scientist, or in the businessman who cannot retire. Such behavior in man is usually accounted for as a search for prestige, but the animal data make this untenable. It seems much more likely that solving problems and running mild risks are inherently rewarding or, in more general terms, that the animal will always act so as to produce an optimal level of excitation. One explanation has been attempted by Hebb (1949b) ; another, into which the facts on the whole would fit very well, is the treatment by Mowrer (1950, 1952) of anxiety reduction as a reward (though here one needs to assume that problem solving is also a cause of anxiety) .

The potential social benefits of a liking for work and (in some essential occupations) for taking risks are sufficiently obvious. This is the plus side of the mammal's liking for excitement. But, as the chimpanzee's behavior shows, excitement may be sought in other less desirable ways. The plain fact is that the primate is only too ready to become a troublemaker when things are dull, and we had better stop comforting ourselves with the accurate but insufficient statement that man has no instinct to make war. He also needs no special coaching to discover a taste for "adventure," and some of his adventures may be socially disastrous. Fear or dislike of the strange is not innate, since it depends on certain prior experiences, yet it still does not have to be taught (Hebb, 1946b) . If, therefore, man is not born with a dislike for those who differ from him in habits or appearance, he can still pick up the dislike with no help or encouragement. Making war is not instinctive; neither, unfortunately, is an aversion to certain forms of excitement that may lead to war.

It appears, therefore, that we stand on very slippery ground when we suppose that social harmony is to be had on a negative basis, by minimizing frustration for the growing child and not teaching him bad ideas about others. Suppose we do give the child as much freedom as we can; suppose the adult does see that he is not being injured or threatened by other groups — is there not still a strong presumption that more will be necessary? That the child must be taught somehow to *like* others different from himself, to get *pleasure* from knowing more about them and learning about their unusual ways? In addition to correcting economic injustice, may it not be necessary also to develop, systematically, socially harmless sources of excitation? Poker and bridge and pennant races, and bowling alleys and detective stories, may have more to do with our social stability than we think. Even today, an economically successful society may require circuses, in some form, as well as bread.

Economic factors obviously play a large part in social conflict; it is only too evident that children can learn hostile attitudes from their elders. We must recognize and control these factors for a sound society. But it is equally essential to recognize that there is no scientific basis for asserting that these are the only factors underlying war and race prejudice. We may even doubt that they are ever effective by themselves, without the deep-seated tendency to be disturbed emotionally by what is different from ourselves.

A Theoretical Approach to Human Nature and Society

We have made some effort in the preceding pages to be economical of theory, using no more than was needed to make a connected story. If we have nevertheless been speculative, we must now become more so. Our purpose in this final section is to offer a theoretical view concerning human emotions which, we believe, casts light on the evolution and structure of society.

Doing so is not outside the scope of the present chapter. The views we set forth derive from the study of animal emotions; and we may remind the reader, if need be, that comparative psychology includes man, regarding him simply as the highest in the animal series. Human culture is unique, but this only makes it the more interesting as a comparative phenomenon. From the comparative point of view it constitutes a special problem.

Man is the most intelligent animal; and the most striking feature of his behavior, comparatively, is an increasing control of the physical environment and the accumulation of "wealth" and wealth-producing skills. But almost equally striking is the progressive modification that has gone on in other aspects of human behavior: broadly speaking, a moral and political development. The economic growth can be treated as the result of a search for comfort by an exceptionally intelligent animal; a unique power of communication, particularly, allows for progressive problem solving. But the cumulative change in other respects: can this be ascribed to the same source?

Can we derive man's social evolution from his exceptional intellectual characteristics? Perhaps we would also allow him some altruism, on occasion, and we must certainly recognize that his powers of learning and communication will perpetuate some mistakes, so that if we see man as intellectually distinctive only we need not imply that all his social arrangements are intelligent ones. Even after making these allowances, however, we may ask whether such a simple model of the origins of society would not be significantly improved by recognizing that man also has exceptional emotional characteristics.

We have already proposed that the mammal seeks excitement when things are dull. This by itself makes an important change in the theory of "economic man." That theory can hold only in an impoverished society (in which the level of frustration is already above the optimum) and *not* in an economically successful one. With a general increase in wealth and security, the risky venture may be preferred to the sure thing, the interesting occupation to one that pays well.

But let us now go further. Evidence from species comparison suggests that emotional susceptibility increases with intellectual capacity. Man is the most emotional as well as the most rational animal. But this susceptibility is partly self-concealing: its possessor tends to seek the environment in which a too-strong emotion is least likely to occur, the one in which disturbance is nearer the optimal level. In man, this makes for the establishment of "civilized" societies, the chief characteristic of which (at least until recently) is not that they improve the economic lot of the average member, but that they provide an environment in which the frequency of acute fear, disgust, anger, and jealousy is decreased. The further a society has advanced along this path, the less subject to strong emotion its members must appear. Violent emotional outbursts vanish, the well-bred man becomes the norm, and psychologists who seldom see the strong, acute disturbance (except in young children) begin to think of adult man as typically unemotional. But the susceptibility to the disturbance may still be there and, although concealed, have a major influence on social organization.

EMOTIONALITY AND INTELLECT

Such views and their comparative origin will be more intelligible if we see how they were arrived at, in the course of constructing a theory presented elsewhere (Hebb, 1949b).

> The first aim was relatively modest: to find some way of bringing together the facts of learning, perception, and set or attitude. This, it seemed, might be feasible if one could treat the thought process as a "phase sequence," as a series of events in the brain each of which might be aroused sensorily or centrally, or both.
>
> The principal question was as to the nature of these events in the brain. Existing anatomical evidence, and some of the psychological evidence, led to the conception that each of the series of activities would be the activity of an "assembly" of neural cells. These assemblies would have to be established before overt learning could occur, which seemed to deny the truth of the theory, until von Senden's (1932) evidence was recalled. The second main difficulty entailed by such ideas, however, is the one that concerns us more. The assembly as postulated would be rather fragile in its function, easily disturbed or disrupted by unusual sensory conditions. The theory thus arrived at the improbable conclusion that a strange environment would interfere with thinking, or might disorganize it completely.
>
> Then came a freeing hypothesis: suppose the interference with thought to be what we already know as *emotion,* and that it is usually mild. The neurophysiological speculations would permit the disturbance to occur in varying degrees, so let us assume that in mild degree it would not disrupt the phase sequence but only slow it. Slowing the phase sequence would mean that thought would continue, with the same general content, for a longer time. The thing thought about (or perceived) would be by definition more "interesting," more capable of arresting and holding attention. A mild disturbance might thus be positively motivating; and a strong disturbance, even though of the same kind, have opposite effects. Since (as we have seen) the behavioral evidence shows that there are such paradoxical effects of emotional stimulation, these neurophysiological notions seemed to have theoretical possibilities even though they still left much to be accounted for.
>
> If such ideas were to be taken seriously, however, they had a further implication which led, rather deviously, to the present problem. If emotion is in some sense a breakdown of the thought process, it seemed that emotionality would become more marked when thought is more complex. The more elaborate the machinery, the more ways its operation can be disturbed. It seemed therefore that development of a higher level of intelligence would mean an increased vulnerability to emotional disturbance.

When the evidence was examined, some support for this idea appeared. The chimpanzee seems to be more subject to emotional disturbance than the dog, and the dog more so than the rat. There were also hints of an ontogenetic correlation, of increased emotionality with intellectual development from infancy to maturity. Now we come to our problem: adult man should be still more emotional than the very emotional chimpanzee and more emotional than the three-year-old child, and he does not seem to be.

But certain aspects of man's social behavior were very suggestive when they were examined from this point of view, and led us to the conclusion that the predicted susceptibility exists, that it is a major influence on social evolution, and that it tends to produce societies such as ours in which it can hardly be detected.

This view depends in part on the phylogenetic correlation. Let us therefore consider it first. There are certain primitive causes of avoidance and aggression that seem to work equally well at all levels of mammalian development: pain stimuli, sudden loud noise, sudden loss of support, restriction of movement by the animal that is used to freedom — all have immediate effects that are comparable in laboratory rat and man. But as we go from rat to man there is a progressive increase in the range of effective stimuli to avoidance and aggression, and in the duration and complexity of the response.

In the rat, for example, there is little need of such a term as "anger" for describing the animal's behavior. A rat is aggressive or he is not, and the aggression has about the same pattern in different circumstances. The same seems true of the dog, although occasionally he shows something that may be homologous with the primate's sulking. But with the chimpanzee, it is essential to distinguish anger from chronic malice if the animal is to be handled safely. The peculiarly human patterns of temper tantrum and sulking occur frequently. The causes of aggression are more varied in the dog than in the rat, and far more varied in the chimpanzee than in the dog. Finally, the period of emotional disturbance following a brief stimulation also increases from rat to dog to chimpanzee (the chimpanzee Fifi, for example, sulked for three weeks over not getting a cup of milk; first showing outright anger, then refusing to accept milk from anyone for a day or so, and continuing for three weeks to refuse it from the one who had denied it to her).

The phylogenetic development of fear susceptibilities is even clearer. In higher mammals' irrational fears, the motionless object may be nearly as much avoided as the moving one. For a lower form like the rat, avoidance of a strange motionless object may occur (Hudson, 1950) but it certainly is not a prominent part of the behavior. The dog falls part way between rat and chimpanzee. Melzack (1952) has described fears in the dog which are obviously like those of the chimpanzee, although the disturbance seems less intense. As for man's nearest relatives, the apes, it would take a page or more to list the specific objects and situations that are known to have caused *persistent* fear in the chimpanzee, as reported in the diaries of the Yerkes Laboratories and in systematic experiments (Köhler, 1927; Yerkes and Yerkes, 1936; McCulloch and Haslerud, 1939; Hebb, 1946b).

When one considers also his manifold angers and hostilities, the chimpanzee at first sight is a bundle of emotional disturbances. This is not really so, of course. Most of the time the chimpanzee is no more angry or afraid than his human caretaker, and frequently less. The thing that one fears another does not. Some animals are "nervous" or short-tempered, but others are placid. Our point here, however, is that as one goes from rat to dog to chimpanzee one finds an increasing variety in the causes of emotional disturbance, an increasing variety of manifestation, and an increasing duration following brief stimulation, all of which is consistent with the idea that susceptibility to emotional disturbance increases with intellectual development.

To this phylogenetic correlation may be added an ontogenetic one. In the observations with anesthetized animals as test objects (described in the preceding section) the frequency and degree of excitation aroused in the half-grown chimpanzees was markedly less than in the full-grown. The clay model of a head hardly got a second glance from the one- and two-year-olds; it absorbed the attention of the five- and six-year-olds but produced no avoidance; in most of the adults it produced strong fear, with some avoidance in all of them. Jacobsen, Jacobsen, and Yoshioka (1932) and McCulloch and Haslerud (1939) have described the increasing emotional responses of the chimpanzee in the first years of life. For man, Jersild and Holmes (1935) have shown a similar increase up to the age of five; they conclude that the trend is thereafter reversed, but we shall try to show that this may indicate, not a reduction of emotional susceptibility after five years, but the success of our society in protecting us from most of the situations that would produce fear in older children and adults.

Jones and Jones (1928) demonstrated that fear of snakes, at least, is much stronger and more frequent in young adults than in children, none of their subjects having had prior exposure to snakes. This was a cross-sectional study, and a longitudinal observation may be added. Two young children, two years apart in age, were encouraged by a parent to handle snakes and liked to do so whenever they saw and could capture one. But the older at the age of eight, the younger at seven, began to show not a fear but a dislike of touching a snake, and this repugnance subsequently became strong in both, still definitely without fear and definitely without any associated injury, in a part of the country where poisonous snakes are never found.

Even without the special theoretical views from which the discussion took off, such phylogenetic and ontogenetic correlations between increasing intellect and increasing emotionality would suggest that thought and emotion are intimately, essentially related. There must be doubt concerning any treatment of emotion as a state or process independent of intellectual processes, and having a separate seat in the nervous system. Brainstem theories of emotion, as developed originally by Head and by Cannon and Bard or more recently by Lindsley (1951), do not seem to give enough weight to the evident cortical elements in the fears and angers of the higher animal. These appear in the kind of perception that can cause emotion in the higher animal (including the role of fantasy reported by the human subject), as well as in the correlation with intellectual level that has just been discussed.

EMOTION AND SOCIETY

Such theoretical considerations lead one to expect that adult man will be more subject to emotional disturbance than the young child or subhuman animal. But such an expectation, surely, is nonsensical. At the age of three to four years the child is apt to be plagued by imaginative fears (Jersild and Holmes, 1935), and to plague his parents with his strong dislikes and outbursts of temper. The adult chimpanzee must be treated, in effect, as a wild animal, and as yet no one has been able to domesticate him. Surely it is adult man who is *least* subject to emotional disturbance. The chimpanzee, it seems, is explosive, unreliable, fearful, dangerous; man (normal man, normally reared) is ordinarily quiet, not given to sudden violent attack on others without warning, not forever in fear of this or that harmless object.

Perhaps so, but let us look closer.

First, in most primitive societies it is clear that man has generally found himself ringed around by malignant ghosts and devils, with beneficent spirits in the minority, and has found it constantly necessary to spend time, effort, and wealth to propitiate even the friendly ones. Fellow members of the tribe commonly possess evil powers; and the one who has the evil eye is feared and hated, or killed, according to circumstances. For any one of the members to be greatly different from the others in appearance, habits, skills, or tastes would be more than apt to cost him his life. As for an outsider, even of the same race, language, and culture — *any* foreigner is distrusted and feared. He may be a convenient source of trophies (a scalp or an embalmed head to show that one is a self-respecting human adult), or alternatively may be captured and brought back whole for torture or ceremonial sacrifice for the edification of the whole tribe.

These are well-established facts of human behavior, and not rare exceptions either. However, if they are cited, and if one cites also the concentration and slave-labor camps of our own generation, to show that man *is* to be thought of as a wild animal, emotionally unstable and often vicious, the answer one will receive is that these things may happen if people are taught as children to behave in these ways, but that fear, hatred, and cruelty are not inherent in man's nature.

If so, one might wonder at the coincidences that allowed the same teaching to originate in so many parts of the globe.

However, we should make it clear at once that we do not argue that these attitudes are inborn. We do argue that man by birth and maturation has a tremendous emotional susceptibility on the one hand and some need for excitement on the other; that if these are handled carefully in the growing child, with a highly specialized environment, the results may be most admirable; but that these qualities by themselves, with no teaching from others or in spite of ineffective teaching, may by reaction to other environments produce socially disastrous results.

As far as we have gone, therefore, and thinking of our own suburban milieux as providing exceptional rather than typical examples of human social behavior, we might be justified in considering man not a tame animal but inherently dangerous, to be domesticated only with great pains. Picking a human society at random from all those that have existed, one might say that the risk taken by a stranger entering it — a member of another society, also chosen at random — would not be less than that taken by a chimpanzee or a wolf encountering a group of his fellows. It is not wholly fantastic, consequently, to suggest that man is more emotionally excited by what is strange than lower animals are, except after special training.

Perhaps we can go further. Even in ourselves, the civilized, amiable, and admirable part of mankind, well brought up and not constantly in a state of fear, there are signs that this urbanity depends as much on our successfully avoiding disturbing stimulation as on a lowered sensitivity. As already suggested, the capacity for emotional breakdown may be self-concealing, leading the animal to find or create an environment in which the stimuli to excessive emotional response are at a minimum. So effective is our society in this respect that its members — *especially* the well-to-do and educated ones — may not even guess at some of their own potentialities. One usually thinks of education, in the broad sense, as producing a resourceful, emotionally stable adult, without respect to the environment in which these traits are to appear. To some extent this may be true. But education can be seen as being also the means of establishing a protective social environment in which emotional stability is possible. Perhaps it strengthens the individual against unreasonable fears and rages, but it certainly produces a uniformity of appearance and behavior which reduces the frequency with which the individual member of the society encounters the causes of such emotion. On *this* view, the susceptibility to emotional disturbance may not be decreased. It may in fact be increased. The protective cocoon of uniformity, in personal appearance, manners, and social activity generally, will make small deviations from custom appear increasingly strange and thus (if the general thesis is sound) increasingly intolerable. The inevitable small deviations from custom will bulk increasingly large, and the members of the society, finding themselves tolerating trivial deviations well, will continue to think of themselves as socially adaptable.

The educated man usually prides himself on his broad-mindedness, his judgment of persons by their essential qualities, and his disregard of the superficial. The college student is, if anything, even more confident of his own contempt for appearances. Now suppose that someone whose appearance, only, departed significantly from the usual pattern applied for a job as a university instructor. Suppose it is a male who takes the highly reasonable view that women's clothing is more comfortable in the North American summer?, or a female who decides that long hair is a nuisance and shaves her head bare. Or suppose it is a war casualty whose face is really disfigured by scar tissue, without ears, nose, eyebrows, or orbital ridges, but whose senses, intellect, and speech are unimpaired: what chance would any of these have to get a job? If he got the job, what chance would there be that students would sit in a classroom and listen to him, no matter how wise he were? None of these are major deviations from the usual human appearance. Merely asking these questions of an undergraduate class allows one to observe a curious mixture of uneasiness and incredulity. The degree of uniformity in the social environment is such that one has difficulty even thinking of a really different one. The trivial changes of fashion from one year to the next look big to us, and we flatter ourselves that a man's appearance does not fundamentally affect our behavior toward him.

Similarly, the adult may think that he is not disturbed by strange places, darkness, or solitude, although young children are. As our lives are usually arranged, this belief is easy to maintain. It is more difficult if one has tried being separated from one's companions in the deep woods at night; reportedly, it is also difficult for those who have suffered solitary confinement.

All this, it seems, casts grave doubt on the idea that man has intrinsically a greater emotional stability than the chimpanzee or other animals. The real difference may be in the protective environment man has created for himself, which he so takes for granted that he becomes unable to see how his stability is achieved, and how his behavior would look to someone not used to it.

More clearly than anywhere else, this is seen in taboo phenomena. The travelers of the eighteenth and nineteenth centuries were completely astonished by taboo — something new, quite foreign to their own lives. Scholars now recognize that taboo has always existed in civilized cultures, but even today it seems not to be recognized how complete and typical its manifestations are, nor how far-reaching in social effect.

Taboo is relevant to the present discussion in several respects. It was drawn to our attention, in the first place, by observing the chimpanzee's response to a chimpanzee or human head (described earlier). This reminded us of the equally puzzling human response to the sight of a major operation or autopsy, so the further observations were made with an apparently dead (i.e., anesthetized) chimpanzee as test object. The ambivalence of the behavior this aroused then recalled Freud's (1919) emphasis on the puzzling ambivalence of some primitive taboos: the object may be simultaneously regarded with reverence, and as dangerous or contaminating. It seems to be thought that in more enlightened societies such anomalies have disappeared. Any remnants of taboo in our own society are thought of simply as "moral prohibitions." Our taboos are thus regarded as purely negative, and more or less based on rational attitudes. But with the chimpanzee's behavior as a starting point, and with the question in mind whether man is really as free of susceptibility to irrational emotion as he seems, we found something different.

Consider first a relatively mild taboo, on the word "God." Although the taboo has become much less strong in recent years, its ambivalence is still beautifully clear. The word (a) has a reverent use, and (b) is also on occasion described literally as a "dirty word," for using which one of the writers some time ago (about 1910) had his mouth washed out with soap. The lexicographer need not be puzzled to find modern examples to illustrate the double meaning of the Latin *sacer*.

A much more powerful taboo on the dead human body is equally ambivalent. Be reverent to the dead, speak in whispers and make no jokes at the funeral service, and, if you are male, remove your hat. There are still remnants of the Middle-Ages and Renaissance prohibition against cutting up the human body, shown in the difficulty of getting bodies for medical-school teaching or permission for autopsy. Dissection violates the sanctity of the dead. But in opposition to such reverence, the corpse is also contaminating, unpleasant, disgusting; disgust may be expressed, for example, at a report that relatives have kissed the dead man after preparation for burial. In general, members of this society are rarely exposed to dead bodies, except in war or disaster, and they tend to be ashamed of their emotional sensitivity thereto, so it is hard to produce formal evidence of the repugnance. An undergraduate class of 198 persons, including some nurses and some veterans, was polled to find out how many had encountered a dead body. Thirty-seven had never seen a corpse in any circumstances; 91 had seen one only after it was prepared for burial — a total of 65 percent who had either not seen a corpse or who had seen one only in the ritualized procedure that rigidly limits exposure to it. These statistics show clearly that society, for some reason, protects its members from something about the dead body, and if there should still be some doubt that this is the emotional disturbance that the body arouses, the reader can get the evidence for himself if he will obtain a human head from the dissecting room and try to carry it home openly in the subway.

Sex taboos are multifarious, but some of them clearly show the ambivalence we are discussing. On the one hand, marriage is holy; on the other, unsanctified sexual activity is — once more — "dirty," contaminating, filthy. We read with amusement and incredulity of the behavior of the primitive people who put to death those who violate a taboo; but the death sentence for rape and the prison sentence for the bigamist or seller of pictures of sexual congress is exactly in the same class, for it does not have to be shown that physical or psychological harm has been done to anyone; only that the taboo is violated.

That there is a double standard of morals is the old way of saying that as a sex object the human female has a stronger taboo than the male. For the ambivalence of taboos, it is significant that the swing from one extreme to the other is greater with the female than the male: her virginity is more sacred than the male's (which is apt to be a source of derision, except when the male in question is under another taboo, of the clergy), and she is more degraded *and dangerous or contaminating to others* when the taboo has been flagrantly broken. Observe also that young men have been known to say, and even on credible report to feel, that they are "not worthy" of some young woman or other; the converse point of view is notably rare.

The only apparent way of accounting for these ambivalences is to suppose that the taboo object is, in the first place, simply a very exciting one; and that this emotional arousal in some circumstances tends to be channeled into one motor outlet (e.g., aggression), in other circumstances another outlet (e.g., fawning or abasement). This implies that there might be individual differences in response to the same object, just as there must be differences in what is exciting to different individuals. It is, of course, evident socially that the growing child left to himself would develop his own set of taboos. "Moral education" has as one of its main objectives to establish uniformity: uniformity of taboo, but more broadly a uniformity of emotional sensitivity.

If all persons could be given the same emotional attitudes, including a fear of acting in a way that arouses emotion too strongly in others, internal social conflict would disappear, if at the same time the society is economically successful and provides its members with comfort and safety. This is the direction, it seems, in which society tends to move, up to the point at which the moral prohibition reduces provocative behavior below the optimal level. This level itself varies, of course, with individuals according to past experience, which means that society will never reach complete uniformity. There will always be a tendency to "test the limits" of the social code; the result in some cases will anger or horrify too many members of society, and the testing will be suppressed. In other cases, however, the tester will elicit a degree of response that is acceptable, and the innovation will be socially perpetuated.

We may see human society, therefore, as conditioned primarily by the avoidance of the too-strong emotional disturbance. In the first place, this may be the action of "economic man": the search for wealth in the form of shelter and safety. Primitive man is likely to have more than enough emotional disturbance imposed on him by insufficient control of his physical environment. But an equally important determinant of early society is a means of controlling man's own behavior, as a source of emotional disturbance. This influence may have definitely uneconomic results, even at an early stage in the development.

If thus we consider man as distinctive not only intellectually but also emotionally, a more adequate picture of human society emerges. Although we have oversimplified, even our account does not make it a simple picture of enforced conformity and avoidance of fear- and rage-provoking stimulation. There are, also, the search for the optimal level of arousal, and individual differences of emotional susceptibility. Our society, as well, has a long way to go before it can be supposed to have reached the maximal degree of conformity. But it does not seem to have been recognized how consistently our society has moved in the direction of arranging that its members act so as not to excite their fellows beyond a certain point. When the strong excitation is not avoided, it is aroused under carefully specified conditions only (marriage, law court, and prizefighting are examples) in which the sexual arousal, fear, or anger aroused in one does not spread uncontrolled to others. Producing a low-level chronic excitation in others, such as fear of losing a job or not passing an examination, is not frowned on; but the hypothetical civilized man, the goal at which approved social education aims, does not, except in the special circumstances referred to, act so as to produce acute, strong sexual arousal, jealousy, anger, or fear. He does not display disgusting objects openly; he is equally careful not to get into situations in which his own emotions make a display of him.

Our social arrangements fall short of this admirable goal, but they have, on the whole, achieved in some segments of this society a remarkable success in avoiding the causes of strong emotional arousal in ordinary life. This result seems quite unpredictable from the assumption that man is in himself a relatively unemotional animal, distinctive mainly by his intellect. It does become intelligible if man has a manifold susceptibility to emotional disturbance; if he tends to avoid actions that lead to strong emotion, and if one of the causes of emotion is the sight of it in others; and if the gradual establishment of conformity makes things become exciting which would otherwise not be so, because of their strangeness.

Man is a rational, unemotional animal as long as there is nothing to disturb his emotions. Also, the causes of strong emotion are few, as we usually consider them; for all we take account of are the causes that *do* operate in a society the main function of which may be to control and limit strong emotion, and we are almost incapable of thinking of an environment that differs from the social and physical one in which we live and which we have so carefully tailored to our needs. The animal evidence shows that the altruism and friendship so prominent in *Homo sapiens* are fundamental to his nature: if we are to learn how to foster such attitudes, we must recognize and provide the kind of protective social environment in which it is possible for them to develop.

References

Allee, W. C., Emerson, A. E., Park, O., Park, T., & Schmidt, K. P. *Principles of animal ecology.* Philadelphia: W. B. Saunders, 1949.

Allport, G. W. The psychologist's frame of reference. *Psychol. Bull.,* 1940, *37,* 1-28.

Beach, F. A. Evolutionary changes in the physiological control of mating behavior in mammals. *Psychol. Rev.,* 1947, *54,* 297-315.

Beach, F. A. *Hormones and behavior.* New York: Hoeber, 1948.

Beach, F. A. The snark was a boojum. *Amer. Psychologist,* 1950, *5,* 115-124.

Berlyne, D. E. Novelty and curiosity as determinants of exploratory behavior. *Brit. J. Psychol.,* 1950, *41,* 68-80.

Bernstein, L. A note on Christie's "Experimental naïveté and experiential naïveté." *Psychol. Bull.,* 1952, *49,* 38-40.

Bierens de Haan, J. A. Animal language in relation to that of man. *Biol. Rev.,* 1929, *4,* 249-268.

Bingham, W. E., & Griffiths, W. J. The effect of different environments during infancy on adult behavior in the rat. *J. comp. physiol. Psychol.,* 1952, *45,* 307-312.

Birch, H. G., & Clark, G. Hormonal modification of social behavior. II. The effects of sex-hormone administration on the social dominance status of the female-castrate chimpanzee. *Psychosom. Med.,* 1946, *8,* 320-331.

Birch, H. G., & Clark, G. Hormonal modification of social behavior. IV. The mechanism of estrogen-induced dominance in chimpanzees. *J. comp. physiol. Psychol.,* 1950, *43,* 181-193.

Carpenter, C. R. Societies of monkeys and apes. *Biol. Sympos.,* 1942, *8,* 177-204.

Clark, G., & Birch, H. G. Hormonal modifications of social behavior. I. The effect of sex-hormone administration on the social status of a male-castrate chimpanzee. *Psychosom. Med.,* 1945, 7, 321-331.

Conant, J. B. *On understanding science: an historical approach.* New Haven: Yale, 1947.

Crawford, M. P. The cooperative solving of problems by young chimpanzees. *Comp. Psychol. Monogr.,* 1937, *14,* No. 68. Pp. 1-88.

Crawford, M. P. The relation between social dominance and the menstrual cycle in female chimpanzees. *J. comp. Psychol.,* 1940, *30,* 483-513.

Crawford, M. P. The cooperative solving by chimpanzees of problems requiring serial responses to color cues. *J. soc. Psychol.,* 1941, *13,* 259-280.

Darling, F. F. *A herd of red deer.* Oxford: Oxford Univ. Press, 1937.

Davis, K. Extreme social isolation of a child. *Amer. J. Sociol.,* 1940, *45,* 554-565.

Davis, K. Final note on a case of extreme isolation. *Amer. J. Sociol.,* 1947, *52,* 432-437.

Dropkin, V. H. Host specificity relations of termite protozoa. *Ecology,* 1941, *22,* 200-202.

Fischel, W. Beiträge zur Soziologie des Haushuhns. *Biol. Zbl.,* 1927, *47,* 678-696.

Ford, C. S., & Beach, F. A. *Patterns of sexual behavior.* New York: Harper (Hoeber), 1951.

Forgays, D. G., & Forgays, J. W. The nature of the effect of free-environmental experience in the rat. *J. comp. physiol. Psychol.,* 1952, *45,* 322-328.

Freud, S. *Totem and taboo.* (Trans. by A. A. Brill.) London: Routledge, 1919.

Goldstein, K. *Human nature in the light of psychopathology.* Cambridge, Mass.: Harvard Univ. Press, 1940.

Goodenough, Florence L. The measurement of mental growth in childhood. In L. Carmichael (Ed.), *Manual of child psychology.* New York: Wiley, 1946. Pp. 450-475.

Greenberg, B. Some relations between territory, social hierarchy, and leadership in the green sunfish *(Lepomis cyanellus). Physiol. Zoöl.,* 1947, *20,* 267-299.

Gregg, R. E. The origin of castes in ants with special reference to *Pheidole morrisi* Forel. *Ecology,* 1942, *23,* 295-308.

Grünthal, E. Zur Frage der Entstehung des Menschenhirns. *Mschr. Psychiat. Neurol.,* 1948, *115,* 129-160.

Hall, C. S., & Whiteman, P. H. The effects of infantile stimulation upon later emotional stability in the mouse. *J. comp. physiol. Psychol.,* 1951, *44,* 61-66.

Harlow, H. F. Levels of integration along the phylogenetic scale: learning aspect. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads.* New York: Harper, 1951. Pp. 121-144.

Harlow, H. F., Harlow, Margaret K., & Meyer, D. R. Learning motivated by a manipulation drive. *J. exp. Psychol.,* 1950, *40,* 228-234.

Hayes, K. J., & Hayes, Catherine. The intellectual development of a home-raised chimpanzee. *Proc. Amer. phil. Soc.,* 1951a, *95,* 105-109.

Hayes, K. J., & Hayes, Catherine. Imitation in a home-raised chimpanzee. *J. comp. physiol. Psychol.,* 1951b, *45,* 450-459.

Hebb, D. O. Emotion in man and animal: an analysis of the intuitive processes of recognition. *Psychol. Rev.,* 1946a, *53,* 88-106.

Hebb, D. O. On the nature of fear. *Psychol. Rev.,* 1946b, *53,* 259-276.

Hebb, D. O. Temperament in chimpanzees: I. Method of analysis. *J. comp. physiol. Psychol.,* 1949a, *42,* 192-206.

Hebb, D. O. *The organization of behavior.* New York: Wiley, 1949b.

Hebb, D. O., & Bindra, D. Scientific writing and the general problem of communication. *Amer. Psychologist,* 1952, 7, 569-573.

Hebb, D. O., & Riesen, A. H. The genesis of irrational fears. *Bull. Canad. psychol. Ass.*, 1943, *3*, 49-50.

Heron, W. T. The inheritance of maze learning ability in rats. *J. comp. Psychol.*, 1935, *19*, 77-89.

Hill, J. C., & Robinson, B. A case of retarded mental development associated with restricted movement in infancy. *Brit. J. med. Psychol.*, 1929, *9*, 268-277.

Hudson, B. B. One-trial learning in the domestic rat. *Genet. Psychol. Monogr.*, 1950, *41*, 99-145.

Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.

Hunt, J. McV. The effects of infant feeding-frustration upon adult hoarding in the albino rat. *J. abnorm. soc. Psychol.*, 1941, *36*, 338-360.

Hunt, J. McV., Schlosberg, H., Solomon, R. L., & Stellar, E. Studies of the effects of infantile experience on adult behavior in rats: I. Effects of infantile feeding-frustration on adult hoarding. *J. comp. physiol. Psychol.*, 1947, *40*, 291-304.

Hymovitch, B. The effects of experiential variations on problem-solving in the rat. *J. comp. physiol. Psychol.*, 1952, *45*, 313-321.

Jacobsen, C. F., Jacobsen, M. M., & Yoshioka, J. G. Development of an infant chimpanzee during her first year. *Comp. Psychol. Monogr.*, 1932, *9*, 1-94.

Jersild, A. T., & Holmes, F. B. *Children's fears.* New York: Teach. Coll. Bur. Publ., 1935.

Jones, E. S. Subjective evaluations of personality. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* Vol. I. New York: Ronald, 1944. Pp. 139-169.

Jones, H. E. Environmental influences on mental development. In L. Carmichael (Ed.), *Manual of child psychology.* New York: Wiley, 1946. Pp. 582-632.

Jones, H. E., & Jones, M. C. A study of fear. *Childhood Educ.*, 1928, *5*, 136-143.

Katz, D. Social psychology and group processes. *Annu. Rev. Psychol.*, 1951, *2*, 137-172.

Katz, I. Behavioral interaction in a herd of Barbary sheep (*Ammotragus Lervia*). *Zoologica*, N. Y., 1949, *34*, 9-18.

Köhler, W. *The mentality of apes.* (2nd Ed.) New York: Harcourt, Brace, 1927.

Levy, D. M. Experiments on the sucking reflex and social behavior in dogs. *Amer. J. Orthopsychiat.*, 1934, *4*, 203-224.

Light, S. F. The determination of the castes of social insects. *Quart. Rev. Biol.*, 1942, *17*, 312-326

Light, S. F. The determination of the castes of social insects. *Quart. Rev. Biol.*, 1943, *18*, 46-63.

Lindsley, D. B. Emotion. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 473-516.

Lorenz, K. A contribution to the comparative sociology of colonial-nesting birds. *Proc. int. orn. Congr., VIII, Amsterdam*, 1934, 207-218.

Lorenz, K. Der Kumpan in der Umwelt des Vogels. *J. Orn., Lpz.*, 1935a, *83*, 137-213.

Lorenz, K. Der Kumpan in der Umwelt des Vogels. *J. Orn., Lpz.*, 1935b, *83*, 289-413.

Lorenz, K. *King Solomon's ring.* London: Methuen, 1952.

McBride, A. F. Meet Mister Porpoise. *Nat. Hist. Mag.*, 1940, *45*, 16-29.

McBride, A. F., & Hebb, D. O. Behavior of the captive bottle-nose dolphin, *Tursiops truncatus. J. comp. physiol. Psychol.*, 1948, *41*, 111-123.

McCulloch, T. L., & Haslerud, G. M. Affective responses of an infant chimpanzee reared in isolation from its kind. *J. comp. Psychol.*, 1939, *28*, 437-445.

McKellar, P. Provocation to anger and the development of attitudes of hostility. *Brit. J. Psychol.*, 1950, *40*, 104-114.

Maier, N. R. F., & Schneirla, T. C. *Principles of animal psychology.* New York: McGraw-Hill, 1935.

Maslow, A. H. The instinctoid nature of basic needs. *J. Pers.*, 1953, in press.

Mason, Marie K. Learning to speak after six and one-half years of silence. *J. Speech Disorders*, 1942, 7, 295-304.

Meehl, P. E., & MacCorquodale, K. A further study of latent learning in the T-maze. *J. comp. physiol. Psychol.*, 1948, *41*, 372-396.

Meehl, P. E., & MacCorquodale, K. Some methodological comments concerning expectancy theory. *Psychol. Rev.*, 1951, *58*, 230-233.

Melzack, R. Irrational fears in the dog. *Canad. J. Psychol.*, 1952, *6*, 141-147.

Miller, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 435-472.

Montgomery, K. C. The relation between exploratory behavior and spontaneous alternation in the white rat. *J. comp. physiol. Psychol.*, 1951, *44*, 582-589.

Montgomery, K. C. Exploratory behavior and its relation to spontaneous alternation in a series of maze exposures. *J. comp. physiol. Psychol.*, 1952, *45*, 50-57.

Morris, C. *Signs, language and behavior.* New York: Prentice-Hall, 1946.

Mowrer, O. H. *Learning theory and personality dynamics.* New York: Ronald, 1950.

Mowrer, O. H. Motivation. *Annu. Rev. Psychol.*, 1952, *3*, 419-438.

Nissen, H. W. A field study of the chimpanzee. *Comp. Psychol. Monogr.*, 1931, *8*, No. 36. Pp. 1-122.

Nissen, H. W. Social behavior in primates. In C. P. Stone (Ed.), *Comparative Psychology* (3rd Ed.). New York: Prentice-Hall, 1951. Pp. 423-457.

Nissen, H. W., Chow, K. L., & Semmes, Josephine Effects of restricted opportunity for tactual, kinesthetic, and manipulative experience on the behavior of a chimpanzee. *Amer. J. Psychol.*, 1951, *64*, 485-507.

Plath, O. E. Insect societies. In C. Murchison (Ed.), *A handbook of social psychology*. Worcester, Mass.: Clark Univ. Press, 1935. Pp. 83-141.

Révész, G. The language of animals. *J. gen. Psychol.*, 1944, *30*, 117-147.

Riesen, A. H. The development of visual perception in man and chimpanzee. *Science*, 1947, *106*, 107-108.

Riesen, A. H. Arrested vision. *Scient. Amer.*, 1950, *183*, 16-19.

Rosenzweig, S. An outline of frustration theory. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. Vol. 1. New York: Ronald, 1944. Pp. 379-388.

Ross, S., Smith, W. I., & Denenberg, V. A preliminary study of individual and group hoarding in the white rat. *J. genet. Psychol.*, 1950, *77*, 123-127.

Sargent, S. S. Reaction to frustration — a critique and hypothesis. *Psychol. Rev.*, 1948, *55*, 108-114.

Schneirla, T. C. Problems in the biopsychology of social organization. *J. abnorm. soc. Psychol.*, 1946, *41*, 385-402.

Schneirla, T. C. A consideration of some conceptual trends in comparative psychology. *Psychol. Bull.*, 1952, *49*, 559-597.

Scott, J. P. Social behavior, organization and leadership in a small flock of domestic sheep. *Comp. Psychol. Monogr.*, 1945, *18*, 1-29.

Sears, R. R. A theoretical framework for personality and social behavior. *Amer. Psychologist*, 1951, *6*, 476-483.

Stewart, Jeannie C., & Scott, J. P. Lack of correlation between leadership and dominance relationships in a herd of goats. *J. comp. physiol. Psychol.*, 1947, *40*, 255-264.

Thompson, W. R. Exploratory behavior as a function of hunger in "bright" and "dull" rats. In press.

Thompson, W. R., & Bindra, D. Motivational and emotional characteristics of "bright" and "dull" rats. *Canad. J. Psychol.*, 1952, *6*, 116-122.

Thompson, W. R., & Heron, W. Exploratory behavior in normal and restricted dogs. In press.

Thorpe, W. H. The concepts of learning and their relation to those of instinct. *Sympos. Soc. exp. Biol.*, 1950, *4*, 387-408.

Tinbergen, N. *The study of instinct*. Oxford: Oxford Univ. Press, 1951.

Tolman, E. C. *Purposive behavior in animals and men*. New York: Century, 1932.

Tryon, R. C. Genetic differences in maze learning ability in rats. *Yearb. nat. Soc. Stud. Educ.*, 1940, *39* (I), 111-119.

Valentine, C. W. The innate bases of fear. *J. genet. Psychol.*, 1930, *37*, 394-419.

Von Frisch, K. *Bees: their vision, chemical senses, and language*. Ithaca: Cornell Univ. Press, 1950.

Von Senden, M. *Raum- und Gestaltauffassung bei operierten Blindgeborenen vor und nach der Operation*. Leipzig: Barth, 1932.

Wasmann, E. *Psychology of ants and of higher animals*. St. Louis: B. Herder, 1905.

Wheeler, W. M. Social life among insects; IV. *Sci. Mon.*, 1922, *25*, 385-404.

Wheeler, W. M. *The social insects, their origin and evolution*. New York: Harcourt, Brace, 1928.

Whiting, J. W. M., & Mowrer, O. H. Habit progression and regression — a laboratory study of some factors relevant to human socialization. *J. comp. Psychol.*, 1943, *36*, 229-253.

Woodworth, R. S. *Psychology*. New York: Holt, 1921.

Yerkes, R. M. Dominance and sexual status in the chimpanzee. *Quart. Rev. Biol.*, 1939, *14*, 115-136.

Yerkes, R. M. Social behavior of chimpanzees: dominance between mates, in relation to sexual status. *J. comp. Psychol.*, 1940, *30*, 147-185.

Yerkes, R. M. Conjugal contrasts among chimpanzees. *J. abnorm. soc. Psychol.*, 1941, *36*, 175-199.

Yerkes, R. M. *Chimpanzees: a laboratory colony*. New Haven: Yale Univ. Press, 1943.

Yerkes, R. M., & Elder, J. H. Oestrus, receptivity, and mating in chimpanzee. *Comp. Psychol. Monogr.*, 1936, *13*, No. 65. Pp. 1-39.

Yerkes, R. M., and Yerkes, A. W. Nature and conditions of avoidance (fear) response in chimpanzee. *J. comp. Psychol.*, 1936, *21*, 53-66.

Zuckerman, S. *The social life of monkeys and apes*. London: Kegan Paul, 1932.

AUTHOR INDEX

Abelson, R. P., 351, 367
Aberle, D. F., 31, 51, 531
Abt, L. E., 499, 519, 521
Ach, N., 44, 51, 199-200, 219
Ackerman, N. W., 174, 176
Adams, D. K., 119, 123, 138, 257
Adler, A., 14, 51, 175
Adler, D. L., 202, 219
Adler, F., 333
Adorno, T. W., 174, 176, 247, 255, 456, 485
Aichorn, A., 174, 176
à Kempis, T., 498
Albert, E., 131, 138
Albig, W., 490, 511, 519
Albrecht, Ruth, 246, 255
Alexander, F., 179
Alexander, G., 235, 255
Allee, W. C., 536-537, 559
Allport, F. H., 23-24, 26, 39-40, 42, 47, 49-51, 58, 69, 79, 85, 87, 126, 138, 341, 366
Allport, G. W., 3, 13-14, 18, 23, 45, 51-52, 76, 87, 93-94, 115-116, 119-120, 122, 129, 133, 138, 145, 173, 224, 237-238, 244, 252, 255, 286-287, 335, 366, 376, 403, 432, 491, 499, 510, 515, 519, 545, 559
Alper, Thelma G., 201, 219
Alpert, A., 108, 138
Ames, A., Jr., 97-98, 113, 138
Ammons, Carol H., 473, 485
Ammons, R. B., 473, 485
Anderson, Gladys L., 499, 519
Anderson, H. H., 208, 219, 392-393, 399, 403, 499, 519
Anderson, T. W., 361, 366
Andrews, T. G., 500, 519
Andreyev, A. A., 499
Angell, R., 178, 485
Angyal, A., 93, 95, 122, 138, 194, 219
Anrep, G. V., 89
Archer, W., 233, 236, 255
Aristippus, 10
Aristotle, 3, 6, 17, 51, 144, 498, 521
Armitt, Fannie M., 287
Arnheim, R., 506, 509, 511, 515, 519
Arnold, Magda B., 237, 255
Arnold, M., 497
Arnold, T., 17, 51
Arrington, R. E., 371, 404
Arsenian, Jean M., 209, 219
Asch, S. E., 18, 23-24, 28-29, 40, 49-51, 92, 106, 113, 119, 122-123, 138, 141, 173, 176, 182, 195, 211, 219, 238, 255
Ash, P., 277, 288, 494, 519
Asheim, L., 492, 517, 519
Ashley-Montagu, M. F., 18, 20, 51, 430, 444
Attneave, F., 260, 264, 287, 348, 366
Auhagen, F. E., 502

Back, K., 181, 215-216, 219-220, 418, 430, 445, 462, 485
Backer, R., 89
Bacon, F., 17, 51
Bagehot, W., 21-22, 24, 51
Baier, D. E., 439, 444
Bailyn, Lotte, 289
Bain, R., 405, 444
Baker, B., 246, 255
Baker, H., 6, 51
Baker, K., 511, 515, 519
Baker, W. B., 429, 446
Baldwin, A. L., 500, 519
Baldwin, J. M., 22, 24, 51, 223, 238, 241-242, 255
Bales, R. F., 172, 174, 176, 179, 374, 376-381, 390, 394, 403, 495, 501, 511, 519
Bard, P., 555
Barker, R. G., 87, 107, 142, 182-183, 193-194, 198, 203, 207, 219, 221-222, 388-389, 403-404, 408, 444, 484-485
Barnard, C. I., 182
Barnes, H. E., 5, 52
Baron, D., 434, 444
Barr, F. E., 429
Barron, M. L., 517, 519
Barry, H., III, 523, 531
Bartlett, D., 289
Bartlett, F. C., 35, 38, 44, 52, 110, 138, 468, 485
Bartlett, K. G., 511, 519
Bartlett, M. S., 275, 287
Bartlett, R. J., 336, 366
Basedow, H., 235, 255
Bassett, R. E., 408, 444
Bastian, A., 35, 52
Bateson, G., 99, 111, 138, 255, 396, 403
Battin, I. L., 308, 333
Bauer, Alice H., 463, 467, 485
Bauer, R. A., 34, 48, 52
Baune, Helen B., 434, 438, 445
Bavelas, A., 197, 199, 202, 213-214, 218-219
Bayes, T., 361
Bayley, Nancy, 241, 255
Bayne, Martha C., 511, 515, 519
Beach, F. A., 523, 530-531, 534, 544, 559
Beaglehole, E., 171, 176
Beard, C. A., 493
Beazley, Virginia (*See* Hertzler, Virginia Beazley)
Bechterew, V. M., 26, 47, 52
Becker, C. L., 493
Becker, H., 5, 20, 52
Beegle, J. A., 412, 428, 430, 438, 445-446

Beeley, A. L., 49
Bell, G. B., 422, 424, 444
Bell, H. M., 433
Bellak, L., 499, 519, 521
Bender, I. E., 237, 247, 256
Benedict, Ruth, 228, 255-256
Benne, K. D., 232, 256
Bennett, Edith B., 217, 219
Bennett, J. F., 337, 354, 366
Bentham, J., 5, 10-13, 16, 52, 53, 55
Berelson, B., 488-489, 494, 496, 498-499, 501-502, 506, 508-512, 515, 517, 519, 522
Bergman, D. V., 385, 403
Bergmann, G., 58, 78, 80, 87, 336, 366
Beria, L., 499
Berlyne, D. E., 551, 559
Bernard, L. L., 5, 22, 41, 49-50, 52, 172, 176
Bernberg, R. E., 342, 366, 369, 461, 487
Bernfeld, S., 144
Bernheim, H., 25-27, 52
Bernreuter, R. G., 433
Bernstein, L., 534, 559
Bettelheim, B., 495, 519
Beum, C. O., 417, 419, 444
Bierens de Haan, J. A., 538, 559
Bindra, D., 532-533, 539, 559, 561
Binet, A., 26-27, 46, 52
Bingham, W. E., 533, 559
Bingham, W. V., 465, 485
Bion, W. R., 174, 176
Birch, H. G., 108, 124, 138, 544, 559
Bird, C., 49
Bismark, O. von, 179
Bitterman, M. E., 124, 138
Bjerstedt, A., 415, 444
Blackwood, Beatrice, 524, 531
Blake, R. R., 138-139
Blanton, Margaret, 23-24, 52
Blanton, S., 23-24, 52
Bliss, C. I., 326
Blum, G. S., 170, 176
Blumenstock, Dorothy, 507, 520
Blythe, Irene T., 493, 519
Boas, F., 170, 172, 528, 531
Bock, R. D., 418-419, 444
Bodine, A. J., 355, 367
Bogardus, E. S., 49, 52, 353, 366
Boguslavsky, G. W., 76, 87
Boldyreff, J. W., 47, 52
Bonney, M. E., 422, 428, 429-433, 437, 444
Borgatta, E. F., 355-356, 366, 369, 405, 412, 421, 444
Boring, E. G., 6, 25, 36, 48, 52, 92-93, 99, 102, 108, 124, 138, 143-144, 176
Bosanquet, B., 34, 52
Bousquet, G. H., 17, 52
Bowlby, J., 174, 176
Bowman, C. C., 236, 256
Braid, J., 25, 27, 52
Brandt, H. F., 208, 219
Bray, D. W., 342, 366
Bretsch, H. S., 434, 446
Brett, G. S., 6, 34, 52
Bridgman, P. W., 91-92, 102, 108, 138
Brierley, Marjorie, 175-176
Brill, A. A., 559
Britt, S. H., 30, 49, 52
Brogden, W. J., 59, 87
Bronfenbrenner, U., 57, 66, 87, 412, 415-416, 420, 428, 444
Brontë, Charlotte, 500, 521
Brontë, Emily, 500, 521
Brown, C. W., 376, 404
Brown, J. C., 228, 256
Brown, J. F., 49-50, 109, 138, 173, 176, 182
Brown, J. S., 65-68, 70, 80-81, 83, 87
Brown, L. G., 49-50
Brown, T., 26, 52
Brown, W., 339, 392
Browne, C. G., 438, 444
Brownfain, J. J., 256
Brücke, E., 144
Brundage, E. G., 417, 444
Bruner, J. S., 59-60, 87, 110, 113, 117, 138, 141, 491, 496, 511, 516, 519
Brunswik, E., 100, 103-104, 132-133, 138, 142, 184, 219, 228, 230, 256
Brush, F. R., 83, 87
Buchanan, R. E., 515, 519
Buchman, M., 141
Bugelski, B. R., 174, 179, 285, 287
Bugental, J. F. T., 244, 256
Bühler, Charlotte H., 132, 138, 371, 403
Bühler, K., 126, 138
Bulganin, N. A., 499
Bunker, H. A., 178
Burchard, W., 252, 256
Burgess, E. W., 55, 171-172, 176, 224, 238, 257
Burt, H. E., 505, 520
Bush, R. R., 65, 73, 77, 83, 87, 89, 289
Buytendijk, F. J. J., 123, 138
Byrd, E., 422-423, 444
Byron, H. J., 20

Cahnman, W., 515, 519
Cameron, N., 224-227, 233, 235-236, 252, 256
Campbell, A. A., 458, 485
Campbell, D. T., 335-336, 342, 366, 460-462, 478, 485, 487
Cannell, C. F., 468, 478, 485
Cannon, W. B., 114-115, 138, 235, 256, 555
Cantril, H., 27, 30, 52, 110, 141, 259, 287, 357, 366-367, 452, 456, 476-477, 485
Caplan, G., 451
Caritat, M. J. A. N. (*See* Condorcet)
Carlson, A. J., 234, 256
Carmichael, L., 221, 486, 559-560

Carpenter, C. R., 538, 541, 559
Carter, H. D., 228, 256
Carter, L. F., 373, 382-384, 394, 403
Cartwright, D., 207, 209, 214, 219-220
Casey, R. D., 520
Cassirer, E., 99, 125-127, 130, 138, 187, 220
Catlin, G., 6, 52
Cattell, J., 140
Cattell, J. McK., 144
Cervinka, V., 419, 444
Chabot, J., 419, 444
Chall, Jeanne, 497, 508, 519
Chapin, F. S., 281, 287, 405, 410-412, 444-445
Chapple, E. D., 384-385, 394-395, 403
Charcot, J. M., 22, 25-27, 30, 46, 52
Chasdi, Eleanor Hollenberg, 82-83, 88, 475, 485, 531
Chase, S., 21, 52
Chave, E. J., 260, 288, 341, 344, 347, 350, 365, 369
Chein, I., 97, 138, 187, 193, 213, 220
Child, C. M., 132, 138
Child, I. L., 58, 71, 83-84, 86-87, 90, 507-508, 519, 523-526, 528-531
Childs, H. L., 520
Chow, K. L., 534, 560
Christiansen, H. F., 281-282, 287
Christie, R., 559
Cisin, I. H., 477, 485
Claparède, E., 112, 138
Clark, G., 544, 559
Clark, K. E., 357, 365-367
Clark, R. A., 82, 87
Clarke, H. M., 44, 52
Cline, V., 246-247, 256
Clyde, R. J., 447
Coch, L., 209, 213, 220
Cochran, W. G., 263, 268, 273-275, 284-285, 287, 328, 333
Coffey, H. S., 389, 403-404
Cohen, A. K., 31, 51
Cologne, Rose, 441, 445
Combs, A. W., 101-102, 122, 141
Comte, A., 3-4, 6-9, 15, 48, 52, 54
Conant, J. B., 532, 559
Condorcet, M. de, 16, 52
Conn, J. H., 473, 485
Cook, S. W., 53, 192, 213, 220, 447, 456, 485-486
Cooley, C. H., 23-24, 52, 223, 238, 242
Coombs, C. H., 335, 337, 343, 358, 366
Cottrell, L. S., Jr., 5, 52, 86-88, 182, 224-225, 227, 243, 247, 256
Coughlin, C. E., 496, 521
Coules, J., 216, 222, 398, 404
Cousins, A. N., 252, 256
Coutu, W., 182, 232, 256
Covner, B. J., 468, 485
Cox, Gertrude M., 263, 268, 273-275, 284-285, 287
Crane, S., 497
Crawford, M. P., 540, 544, 559
Crawley, E., 235, 256
Crespi, L. P., 471, 485
Cressey, D. F., 248, 258
Criswell, Joan H., 405, 411, 415-416, 419, 422, 425-426, 428-430, 443-445
Cronbach, L. J., 259, 264, 287, 338-339, 366, 397, 403
Crossley, Helen M., 480, 486
Crown, S., 365-366
Crutchfield, R. S., 18, 27, 49-50, 54, 91, 94, 107, 140, 173, 179, 182, 211, 221, 238, 257, 335, 362, 367, 452, 456, 485
Cureton, E. E., 333
Curran, F. J., 440, 446
Curti, M. E., 6, 52

Dale, E., 497, 506, 508, 519
Danielsson, B., 423, 430, 445
Darling, F. F., 541, 559
Darwin, C., 16, 41, 52, 143
Dashiell, J. F., 47, 52
Davidson, D. M., Jr., 409, 441, 446
daVinci, L., 175, 177
Davis, A., 86, 88
Davis, A. K., 31, 51
Davis, F. B., 352, 366
Davis, K., 172, 176, 533, 559
DeGrange, McQ., 7, 52
deGrazia, S., 498-499, 502, 508, 511, 517, 519
Delaisi, F., 17, 52
deLange, M., 47, 52
DeLott, J., 469, 487
Dembo, Tamara, 87, 202-203, 208, 219-221
Denenberg, V., 534, 561
Dennis, W. D., 135, 138-139
De Saussure, R., 175-176
Descartes, R., 17, 41
Deutsch, F., 500, 519
Deutsch, K., 188, 220
Deutsch, M., 53, 181, 192, 202, 214-216, 220, 399, 403, 447, 456, 485-486
Deutschberger, P., 431, 441, 445
Devereux, G., 171-172, 176, 235, 256
Dewey, J., 11, 17, 41-42, 48, 51-52, 223
Dewey, R., 73, 88
Dickens, C., 11
Dickson, Lenore, 409, 430, 432, 446
Dickson, W. J., 438, 451, 486
Dietrich, D. H., 468, 487
Dinsmoor, J. A., 79, 88
Dittman, A. T., 385, 403
Dixon, W. J., 285, 287, 289, 333
Dollard, J., 12, 23-24, 42, 52, 55, 58, 62, 66, 70, 73, 76, 81-86, 88-89, 97, 119, 141, 145, 174, 176, 227, 257, 286, 288, 501, 519
Donne, J., 497
Doob, L. W., 21, 30, 49, 52, 54, 69, 82-83, 88, 97, 119, 138, 174, 176, 335, 366

Dostoevsky, F., 500, 520
Droba, D. D., 45, 52, 344, 366
Dropkin, V. H., 537, 559
Dryden, J., 497
DuBois, Cora A., 171, 176, 226, 241, 256
Dudman, J., 361
Duncan, D. B., 304, 333
Duncker, K., 103, 107-109, 138-139
Dunlap, K., 41, 49, 52-53
Durkheim, E., 22, 37-38, 48, 53
Dymond, Rosalind F., 237, 247, 256

Ebel, R. L., 363, 366
Eberhart, J. C., 477-478, 486
Edgerton, H. A., 353, 366
Edwards, A. L., 259-260, 263-264, 268-269, 273-274, 278, 282, 284-287, 289, 291, 303-304, 333, 348, 351, 357, 365-366
Edwards, Daisy S., 415, 445
Ehrenfreund, D., 131, 138
Eichler, R. M., 326, 333
Einstein, A., 182, 220
Eisenhart, C., 289, 322, 334
Eissler, K. R., 174, 176
Elder, J. H., 544, 561
Elinson, J., 477, 485
Ellwood, C. A., 49, 51, 53
Emerson, A. E., 536, 559
Emerson, R. W., 497
Emery, D., 181
Eng, E., 408, 445
Engelmann, G., 5, 53
Engels, F., 11, 17, 54
England, L. R., 463, 485
Entwisle, Doris, 289
Epicurus, 10
Erikson, E. H., 169, 171, 176
Escalona, Sybille K., 207, 220
Espinas, A., 35, 53
Estensen, E. V., 490, 519
Estes, W. K., 73, 77, 88
Ewer, B. C., 49
Eysenck, H. J., 365-366

Faden, Janet M., 510, 515, 519
Faigin, H., 531
Fajans, Sara, 206, 220
Faraday, M., 25, 53, 181
Farber, I. E., 65-68, 70, 80-81, 83, 87
Farberow, N. L., 249, 258
Faris, E., 45, 53, 172, 176, 238
Faris, R. E. L., 49
Farnsworth, P. R., 49
Faunce, D., 430, 438, 445
Fechner, G., 144
Feigl, H., 58, 88, 102, 107, 139
Feldman, J. J., 477-478, 485
Feldstein, M. J., 409, 422, 447
Fenichel, O., 164, 176, 252, 256
Féré, C., 43, 46, 53
Ferguson, A., 336, 366
Ferguson, G. A., 352, 366
Ferguson, L. W., 351, 366
Ferris, C. D., 289, 291, 302, 333
Festinger, L., 182, 207-208, 214-217, 220-221, 356, 366, 418-419, 430, 445, 485
Fichte, J. G., 35
Fischel, W., 541, 559
Fisher, R. A., 48, 261, 268, 270-271, 273-274, 285, 287, 312, 329-331, 333
Fiske, D. W., 342, 367
Fitch, F. B., 88
Fitzpatrick, G. A., 439, 448
Flemming, G., 433, 445
Flesch, R., 497, 508, 519
Flowerman, S. H., 471, 476-478, 487
Foa, U. G., 414, 445
Folsom, J. K., 49, 172, 176
Foote, N. N., 86, 88
Ford, C. S., 232, 256, 523, 525, 528, 530-531, 544, 559
Ford, R. N., 355, 366
Forgays, D. G., 533, 559
Forgays, J. W., 533, 559
Forlano, G., 351, 368, 433, 447
Forsyth, Elaine, 416-420, 432, 445
Foster, C. R., 505, 515, 519
Foster, H. S., 506, 519
Fouillée, A., 38, 53
Fouriezos, N., 381, 404
Frank, J., 175-176
Frank, J. D., 209, 220
Franklin, Muriel, 140, 202, 221
Fredericson, E., 240, 256
Freedman, H., 433, 447
Freedman, M. B., 389-390, 403-404
Freedman, R., 281, 287
Freeman, E., 49
Freeman, M. F., 326-327, 333
Freeman, Willa (*See* Grunes, Willa Freeman)
French, J. R. P., Jr., 66, 88, 185, 209, 213-215, 220
French, R. L., 408, 422, 424, 432, 435, 444-445
French, T. M., 168-169, 176
Frenkel-Brunswik, Else, 176, 485
Freud, Anna, 163-164, 168, 176, 178-179, 251-252, 256
Freud, S., 5, 11-12, 14, 16, 18, 24-26, 28, 31, 35, 41, 44, 53, 58, 77, 82, 115, 143-180, 188, 223, 239, 256, 455, 506, 520, 557, 559
Friedlander, Kate, 174, 178
Friedman, G. A., 174, 179, 523, 528, 531
Friedman, M., 320, 333
Friedman, O., 49
Fries, C. C., 497, 519
Fritzsche, H., 518
Fromm, E., 115, 139, 170, 173-175, 178
Fromme, A., 342, 366
Frosch, J., 176

Fryer, D., 425, 439, 448
Fuller, Elizabeth M., 434, 438, 445
Fuson, W. M., 335, 367

Gabel, Hortense, 517, 519
Gabel, M., 517, 519
Galileo, 109
Gall, F. J., 8
Gallagher, Ruth, 52, 243, 256
Gallup, G., 48, 457
Galton, F., 48, 53, 523
Gardiner, B. B., 461, 485
Garner, W. R., 348, 367
Gault, R. H., 49
Gauss, C. F., 109
Gaylord, R. H., 352, 369
Gehlke, C. E., 38, 53
Gelb, A., 128, 139
George, A., 503
George, H., 9
George III, 492
Gerard, H. B., 181, 216-218, 220
Gerbrands, H., 394, 403
Gerson, J., 498
Getzels, J. W., 460-461, 485
Ghiselli, E. E., 376, 404
Gibb, C. A., 425, 445
Gibson, Eleanor, J., 64, 80, 88
Gibson, J. J., 85, 88, 100, 112-113, 117, 132, 139
Gibson, W. A., 361
Giddings, F. H., 20, 53
Gillin, J. P., 86, 88, 525, 531
Ginsberg, M., 523, 531
Glaser, N., 462, 486
Glixman, A. F., 201, 220
Godwin, W., 16, 53
Goebbels, J., 503, 518
Goethe, J. W., 136
Goldenweiser, A., 35, 38, 53, 531
Goldman, Frieda (*See* Goldman-Eisler, Frieda)
Goldman-Eisler, Frieda, 239, 256
Goldsen, J., 488-489, 510, 514-515, 520
Goldstein, K., 93-95, 111, 114-115, 122, 125-128, 132-134, 139, 141, 535, 559
Goldwyn, R., 435, 446
Gonick, Mollie R., 198, 219
Goodacre, D. M., III, 428, 440, 445
Goodenough, Florence L., 264, 287, 371, 480, 485, 533, 559
Goodenough, W. H., 356, 367
Goodman, C. C., 60, 87
Goodnow, R. E., 426, 429, 445
Gordon, D. A., 452, 456, 485
Gordon, L. V., 362, 367
Gordon, M. H., 329, 331, 333
Gottheil, E., 421-422, 445
Gottlieb, Lillian, 517, 520
Gottschalk, L., 178, 485
Gough, H. G., 227, 232-233, 244-246, 248, 251, 256
Goulden, C. H., 285, 287
Grant, D. A., 284, 287
Gray, W. S., 497, 508, 520
Green, A. W., 172, 178
Green, B. F., 335, 356, 360-361, 367
Green, Helen D., 441, 445
Green, T. H., 34, 53
Greenberg, B., 541, 559
Greenwood, E., 53
Gregg, R. E., 536, 559
Grey, A., 514, 516, 520
Griffiths, W. J., 533, 559
Gronlund, N. E., 437, 445
Gross, N., 459, 463, 485
Grosser, D., 210, 220
Grossmann, Beverly, 429, 431, 433, 445
Grubbs, F. E., 289, 291, 333
Grunes, Willa Freeman, 222
Grünthal, E., 543, 559
Guest, L., 471, 476, 478, 485
Guetzkow, H., 179, 381, 398, 404, 450, 485
Guilford, J. P., 264, 287, 305, 319-320, 333, 345, 347, 367
Gulliksen, H., 264, 287, 336, 339-340, 348, 352-353, 363, 367
Gumplowicz, L., 4
Gurnee, H., 49
Guterman, N., 30, 54, 174, 179
Guthrie, E. R., 58-59, 63, 65, 68, 70, 72-76, 84, 88, 286-287
Guttinger, F., 517, 520
Guttman, L., 260, 287, 335-336, 339, 344, 353-358, 361-362, 364-367, 397-398, 404, 477, 481

Hacker, Helen M., 228, 256
Haddon, A. C., 5, 53
Hake, H. W., 348, 367
Hale, W. B., 498, 520
Hall, C. S., 143, 534, 559
Hall, G. S., 48, 53, 144, 223
Hall, J. K., 178
Hall, M., 88
Hall, W. B., 223
Hallowell, A. I., 171, 178
Hamilton, A., 16
Hamilton, C. H., 309, 334
Hamilton, T., 508, 520
Hamilton, W., 26, 53
Hammond, K. R., 342, 367, 461, 485
Hamsun, K., 500
Hanchett, Gertrude A., 341, 366
Hanfmann, Eugenia, 126, 128, 132, 139
Harary, F., 199, 220, 418, 447
Harding, J., 181, 213, 220
Hardyck, C., 230, 248, 257, 258
Harlow, H. F., 118, 124, 139, 550-551, 559
Harlow, Margaret K., 551, 559

Harriman, P. L., 138
Harris, H., 510, 520
Harris, M., 267, 287
Harsh, C. M., 246, 256
Hart, A., 498, 520
Hart, C. W., 477-478, 485
Hart, H., 4, 48, 53, 494, 503, 505, 520
Hartley, E. L., 49, 174, 179-180, 221, 229, 238, 256, 531
Hartley, F. A., 494, 519
Hartley, Ruth E., 49, 238, 256
Hartmann, G. W., 240, 258
Hartmann, H., 115, 139, 168-169, 178
Hartshorne, H., 353, 368, 410, 445
Haslerud, G. M., 554, 560
Hastorf, A. H., 237, 247, 256
Hatch, D. L., 515, 520
Hatch, Mary A., 515, 520
Havighurst, R. J., 86, 88
Hawley, A. H., 281, 287
Hayakawa, S. I., 17, 53
Hayes, Catherine, 539-540, 559
Hayes, K. J., 539-540, 559
Hays, D. G., 355-356, 361, 366-367, 369
Haythorn, W., 382, 403
Hayworth, D., 498, 520
Hayworth, Rita, 125
Head, H., 110-112, 126-127, 139, 555
Hebb, D. O., 66, 69, 88, 100, 139, 142, 532, 539-540, 543-545, 547, 549, 551-554, 559-560
Hegel, G. W. F., 3, 34-36, 53
Heidbreder, Edna, 127-128, 139
Heider, F., 96-97, 105-106, 119, 139, 182-183, 194, 199, 210-212, 216, 220
Helmholtz, H., 144
Helson, H., 110, 139, 178
Hendrick, I., 168, 178
Heneman, H. G., Jr., 465, 485
Henle, Mary, 202, 220
Henmon, V. A. C., 246
Henry, A. F., 355-356, 367, 369
Henry, J., 171, 178
Henry, Zunia, 171, 178
Herbart, J. F., 25, 36, 53
Herma, H., 506, 520
Heron, W. T., 533-534, 560-561
Herrick, C. J., 125, 139
Hershkowitz, A., 91
Herskovits, M. J., 256
Hertz, P., 181
Hertzler, Virginia Beazley, 428-432, 446
Herzog, Herta, 509, 522
Hevner, Kate, 260, 287, 347, 364, 367
Heyns, R. W., 370, 378, 404
Hilgard, E. R., 59, 65, 76, 88, 92, 95, 100, 116, 118, 120, 124, 139, 170, 178, 208, 220, 244, 257, 465, 485
Hill, H. W., 138, 142
Hill, J. C., 533, 560
Hinckley, E. D., 351, 367
Hitler, A., 4, 30, 34, 492, 520-522
Hobbes, T., 3, 10, 13-16, 32-33, 53, 79
Hobhouse, L. T., 523, 531
Hocking, W. E., 39, 53
Hoel, P. G., 263, 287
Hollenberg, Eleanor (*See* Chasdi, Eleanor Hollenberg)
Hollingworth, H. Leta, 26, 53
Holmes, F. B., 554-555, 560
Holt, E. B., 23-24, 42, 53
Holzberg, J. D., 435, 445
Holzinger, K. J., 418, 444
Homans, G. C., 70, 88, 182
Hoppe, F., 13-14, 53
Horney, Karen, 14, 53, 169-170, 173-174, 178
Horowitz, E. L., 341-342, 367-368
Horowitz, M., 218, 220, 222
Horowitz, M. W., 212, 220
Horowitz, Ruth E., 341-342, 367-368
Horst, P., 353, 367
Horton, C. P., 76, 88
Horton, D., 493, 521, 523, 530, 531
Horvitz, D. G., 267, 287
House, F. N., 5, 53
Hovland, C. I., 59, 61-62, 84, 88, 248, 277, 287, 351, 367, 450, 485
Howes, D. H., 85, 90
Hudson, B. B., 554, 560
Hughes, E. C., 224, 257
Hughes, J. W., 224, 257
Hull, C. L., 26, 53, 58-59, 63-76, 80-81, 83, 85, 88, 90, 92-93, 95-96, 100, 120, 124, 129-131, 139, 145, 169, 188, 286-287, 333, 542, 560
Humber, W. J., 73, 88
Hume, D., 103-104, 532
Humphrey, G., 23-24, 53
Hunt, J. McV., 84, 88-89, 180, 221-222, 239, 257, 422, 425, 431, 445, 534, 560-561
Husain, Suraya Z., 418-419, 444
Husserl, E. G., 101
Hutt, M., 381, 404
Huxley, J., 37, 125, 139
Hyde, R. W., 440, 445
Hyman, H., 469, 471, 477-478, 480, 485-487
Hymovitch, B., 485, 533, 560

Immerglück, L., 141
Infeld, L., 182, 220
Infield, H. F., 441, 445
Irwin, F. W., 286-287
Irwin, Lucietta, 222
Isaacs, Susan, 135, 139

Jack, Lois M., 374, 391-392, 404
Jackson, J. H., 126-127, 139
Jacob, P. E., 512, 520
Jacobs, J. H., 438, 445

Jacobsen, C. F., 554, 560
Jacobsen, M. M., 554, 560
Jacobson, E., 237, 257
Jaensch, E. A., 35, 53
Jahoda, Marie, 53, 174, 176, 181, 192, 220, 447, 456, 485-486
James, W., 22, 25, 34, 37, 41-42, 44, 46, 48, 53, 123, 139, 144, 160, 223, 238-239, 257
Janet, P., 25-26
Janis, I. L., 489, 520
Janowitz, M., 172, 180, 495, 519
Jaspen, N., 277, 288
Jefferson, T., 16, 18, 28-29
Jeffress, L. A., 140
Jenkins, J. G., 428, 440, 446
Jenkins, R. L., 440, 446
Jenkins, W. O., 58, 76, 83, 87-88
Jennings, Helen H., 406, 409, 411, 413-416, 419, 422-424, 427, 434, 437, 439, 443, 445-447
Jersild, A. T., 554-555, 560
John, W. F., 83, 87
Johnson, A. D., 437, 446
Johnson, H. M., 336, 367
Johnson, P. O., 263, 274, 285, 288
Jones, Dorothy B., 506, 509-511, 520
Jones, D. S., 223
Jones, E., 16, 53, 172, 178, 500, 520
Jones, E. S., 545, 560
Jones, H. E., 533, 554, 560
Jones, L. V., 262, 288
Jones, Mary C., 118-119, 139, 554, 560
Jordan, N., 212, 220
Jowett, B., 51, 55
Jucknat, M., 208, 220
Judd, C. H., 35, 39, 53
Jung, C. G., 34, 53, 146, 175

Kaganovich, L. M., 499
Kahane, M., 409, 446
Kahn, L. A., 355, 367
Kahn, R. L., 482, 485
Kamin, L. J., 85, 90
Kamiya, J., 253, 258
Kant, I., 25
Kantor, J. R., 49
Kanzer, M., 500, 520
Kaplan, A., 488-489, 510, 514-515, 520
Kaplan, D., 514, 516, 520
Kaplan, Edith, 132, 136, 142
Kardiner, A., 166, 170-171, 178
Karpf, Fay B., 5, 53, 172, 178
Karsten, Anitra, 202, 220
Kasanin, J., 126, 132, 139
Katona, G., 11, 54, 109, 139, 481, 485
Katz, D., 49, 54, 357, 367, 471, 476-477, 482, 485, 551, 560
Katz, I., 541, 560
Katz, L., 416-420, 425, 432, 445-446
Kearl, B., 497, 520
Kecskemeti, P., 174, 178, 504, 520
Keet, C. D., 385, 404
Keller, Helen, 127, 139
Kelley, H. H., 216, 218, 220, 351, 367, 485
Kelly, E. L., 342, 367
Kempf, E. J., 16, 54
Kempthorne, O., 285, 288
Kendall, M. G., 264, 288, 320, 334
Kendall, Patricia L., 451, 486
Kendler, H. H., 107, 139
Kenney, Kathryn C., 260, 287, 365-366
Kerr, W., 434, 438, 448
Kerstetter, Leona M., 437, 441, 446, 448
Khrushchev, N. S., 499
Kilpatrick, F. P., 260, 287, 357, 366
Kimble, G. A., 61, 88
Kingsbury, Susan M., 494, 520
Kingsley, C., 11, 500, 519
Kinsey, A. C., 334, 451-453, 457, 463, 466, 480, 485
Klein, G. S., 69, 88, 113, 116-117, 140, 143
Klineberg, O., 49
Kluckhohn, C., 5, 12, 35, 37, 54, 86, 89, 170-171, 178, 255, 451, 473-474, 485, 528, 530-531
Kluckhohn, Florence, 475, 486
Koch, S., 58, 88
Koffka, K., 44, 54, 95, 97-98, 101, 103, 110-113, 122, 133, 140, 191, 221, 238, 257, 286, 288
Köhler, W., 23-24, 54, 92, 96-98, 103-104, 108, 113, 122-124, 130-131, 140, 539, 542-543, 549-550, 554, 560
Kolbe, L. E., 353, 366
Komarovsky, Mira, 228, 257
Korsakow, S., 112
Kosygin, A., 499
Kounin, J. S., 201-203, 219, 221
Kracauer, S., 513, 520
Krech, D., 18, 27, 49-50, 54, 69, 88, 91, 94-95, 107, 111, 140, 173-174, 178-179, 182, 211, 221, 238, 257, 335, 362, 367, 532
Kriedt, P. H., 357, 365-367
Kriesberg, M., 516, 520
Kris, E., 143-144, 168-169, 174, 178-179, 491, 500, 506, 508, 512, 520
Kroeber, A. L., 5, 37, 54, 170, 179, 486, 527, 531
Kropotkin, P. A., 18-20, 54
Krout, M. H., 49
Krueger, E. T., 49
Krugman, Dorothy C., 256
Kruskal, W. H., 289, 320, 334
Kubie, L. S., 170, 179
Kuder, G. F., 339, 432
Kuenne, Margaret R., 131, 140
Kuhlen, R. G., 434, 446
Külpe, O., 144
Kurtz, A. K., 277, 288

Lambert, Elisabeth C., 57
Lambert, W. W., 57, 60, 62, 85, 88
Lamoreaux, R. R., 131, 141
Landis, M. H., 505, 520
Lange, C., 160
Lange, L., 43, 54
Lange, N., 43, 54
Langer, Susanne K., 127, 129, 140
Lanzetta, J., 382, 403
LaPiere, R. T., 49, 341, 367
Lashley, K. S., 95, 98, 110-111, 127, 130-131, 140
Lasswell, H. D., 175, 179, 219, 485-486, 490-491, 495-496, 499-505, 507-508, 510, 512, 514, 516, 520-522
Lavoisier, A. L., 532
Lawrence, David H., 500, 521
Lawrence, Douglas H., 61, 88
Lawsing, Margaret, 429, 446
Lazarsfeld, P. F., 335, 343, 355, 359, 361, 367, 416, 455, 486, 490, 495, 519, 521
Lazarus, M., 3, 35-36, 54
Leary, B. E., 497, 508, 520
Leary, T. F., 389, 403-404
Leavitt, H. J., 199, 218, 221, 409, 425, 439, 448
LeBon, G., 22, 25-26, 29-31, 54
Lecky, P., 14, 54, 239, 257
LeDantec, F., 14-15, 17, 54
Lee, A. McC., 496, 512, 521
Lee, Elizabeth B., 496, 512, 521
Lee, I. J., 140
Lee, V., 17, 54
Leeman, C. P., 442-443, 446
Leeper, R. W., 65, 88, 93, 109, 122, 140, 190, 196-197, 199, 221
Leighton, A. H., 21, 54
Leites, N., 174, 178, 488-489, 491, 499, 504-505, 508, 511-512, 517-518, 520-522
Lemann, T. B., 420, 427, 432, 446
Lenin, N., 28-29, 499
Lerner, D., 219, 485-486, 506, 521
Levine, Estelle M., 507-508, 519
Levinson, D. J., 176, 485
Levy, D. M., 170, 179, 473, 486, 534, 560
Levy, M. J., 31, 51
Levy, N., 89
Levy, S., 215-216, 221
Levy-Bruhl, L., 7, 38, 54
Lewin, H. S., 492, 521
Lewin, K., 47-48, 54, 58, 67, 77-78, 87-88, 94, 101, 140, 145, 169, 173-174, 181-193, 195-209, 212-214, 217-219, 221-223, 238, 257, 286, 288
Lewis, D. J., 89
Lewis, Helen B., 122, 140, 202, 221
Lewis, J. H., 234, 257
Lewis, M. M., 242
Lewis, P. M., 510, 520
Lichtenberg, P., 181
Liebeault, A. A., 25, 54
Light, S. F., 536, 560
Likert, R., 260, 287-288, 344, 351-353, 357, 360, 364-366, 368
Lindemann, E., 385, 403
Lindesmith, A. R., 49, 51, 87-88, 126, 140, 257
Lindner, R. N., 499, 521
Lindquist, E. F., 269, 274, 288, 366, 368
Lindsley, D. B., 555, 560
Lindzey, G., 50, 54, 143, 174, 179, 405, 434-435, 446
Link, R. F., 289, 304, 320, 334
Linnaeus, C., 537
Linton, R., 171, 179, 224, 257
Lippitt, R., 48, 54, 209-210, 212-215, 218, 220-222, 370, 374, 387, 394, 396-397, 404, 423, 447
Lippmann, W., 17, 54
Lipps, T., 23-24, 54, 231
Lissner, Kate, 202, 222
Littman, R. A., 95, 140, 183, 222
Locke, H. J., 172, 176
Locke, J., 15, 54
Loevinger, Jane, 339, 357, 368
Loewenstein, R. M., 168-169, 174, 178-179
Lombroso, C., 30
Longmore, T. W., 412, 428, 446
Loomis, A. M., 371, 404
Loomis, C. P., 409, 411-412, 414-417, 421, 425, 428-430, 432, 441, 446-447
Lord, E., 323-324, 334
Lord, F. M., 352, 368
Lorenz, K., 534-535, 541, 560
Lorge, I., 337, 350, 368, 497, 508, 521
Loveland, E. H., 333
Lowell, A. L., 534
Lowenfeld, H., 175, 179
Lowenthal, L., 30, 54, 174, 179, 500, 504, 509, 511, 517, 521
Luce, D. R., 418, 446
Luchins, A. S., 110, 117, 140, 532
Lukasczyk, K., 5, 35, 54
Lumsdaine, A. A., 88, 277, 287, 450, 485
Lundberg, G. A., 405, 407-409, 428-432, 446
Lynip, A. W., 126, 140
Lyons, J., 212, 220

McBride, A. F., 542-543, 560
McCarthy, Dorothea, 135, 140
McClelland, D. C., 70-71, 85, 89, 174, 179, 224, 257, 523, 528, 531
McClelland, F. M., 437, 446
Maccoby, Eleanor E., 449
Maccoby, N., 449, 485
MacCorquodale, K., 335, 368, 550, 560
McCulloch, T. L., 554, 560
McCurdy, H. G., 500, 521
McDougall, W., 5, 9, 12-14, 16-18, 20-25, 27-28, 32, 35, 39-42, 45-46, 48, 50-51, 54, 58, 144-145, 179, 188
McGarvey, H. R., 364, 368
McGinnies, E., 117, 141
McGranahan, D. V., 492, 504, 511, 521

McGregor, D., 209, 222
Macgregor, Frances C., 396, 404
McGuffey, W. H., 490, 519
Mach, E., 92
Machiavelli, N., 15-16, 54
McKellar, P., 550, 560
McKinney, J. C., 414-415, 446
MacLeod, R. B., 18, 54, 58, 89, 93, 101-102, 106-107, 119, 122-123, 140
McNemar, Q., 219, 274, 284-285, 288, 331, 334-335, 368, 479, 486
McReynolds, P., 58, 89
MacStewart, W., 314, 334
Madison, J., 16
Magaret, G. Ann, 208, 220, 226, 233, 256
Mahan, A. T., 493
Mahler, Wera, 202, 222
Mahut, 551
Maier, N. R. F., 21, 54, 66, 89, 109, 131-132, 140, 538, 560
Malenkov, G. M., 499
Malinowski, B., 35, 54, 78, 171-172, 179, 527
Maller, J. B., 410, 445
Mann, F., 482, 486
Mann, H. B., 290, 315-316, 323, 325, 334
Mannheim, K., 11, 17, 54
Mantel, N., 323, 334
Marder, E., 355, 368
Marks, M. R., 262-263, 288
Marrow, A. J., 201, 222
Martin, C. E., 480, 485
Martin, E. D., 31, 54
Marx, K., 11, 17, 34, 54
Marx, M. H., 79, 89, 93, 140
Maslow, A. H., 114-116, 140, 550, 560
Mason, Marie K., 533, 560
Mason, W., 459, 463, 485
Massarik, F., 409, 446
Massey, F. J., Jr., 285, 287
Mather, K., 285, 288
Mathews, Ravenna, 248, 257
Maucorps, P. H., 409, 439, 446
Maurel, H., 53
Maxwell, C., 181-182, 222
May, M. A., 21, 52, 54, 120, 141, 353, 368, 410, 445
Mayer, A., 46, 54
Mazelis, Sarah, 232, 257
Mead, G. H., 22, 24, 54, 86-87, 89, 182, 223-224, 226-227, 232-233, 238, 242-243, 245, 257, 539
Mead, Margaret, 21, 54, 171-172, 179, 255, 396, 403-404, 430, 447, 473, 486
Meehl, P. E., 57, 89, 335, 368, 550, 560
Meirowitz, Beatrice, 382, 403
Meltzer, H., 13, 54
Melzack, R., 551, 554, 560
Mensh, I. N., 106, 141, 432, 445
Merrill, K., 219
Merton, R. K., 224, 249, 257, 451, 454, 486, 507-508, 521
Mesmer, F. A., 25, 54
Metzner, Helen, 482, 486
Meyer, D. R., 551, 559
Meyerson, A., 49
Meyerson, L., 91
Michotte, A., 104-106, 141
Mikoyan, A. I., 499
Miles, Josephine, 490, 497, 508, 512, 521
Mill, J. S., 10-13, 16, 55
Millay, Edna St. V., 497
Miller, D. R., 174, 179, 435, 447
Miller, G. A., 85, 89, 91, 141
Miller, N. E., 12, 23-24, 42, 52, 55, 58-59, 62, 66, 70-71, 73-74, 76, 81-86, 88-89, 97, 119, 124, 141, 145, 174, 176, 179, 206, 222, 227, 257, 286, 288, 542, 550, 560
Miller, S. M., 463, 486
Minturn, L., 85, 89
Mintz, A., 516, 521
Mitrano, A. J., 85, 89
Moede, W., 47, 55
Molotov, V. M., 499
Montgomery, K. C., 551, 560
Mood, A. M., 263, 267, 287-289, 314, 322, 333-334
Mooney, R. L., 434
Moore, B. V., 465, 485
Moore, H. T., 47, 55
Moore, O. L., 58, 89
Moreno, J. L., 48, 55, 227, 230, 236, 253, 257, 405, 407-411, 413-414, 416-417, 422-423, 425-430, 436-438, 441-443, 447
Morgan, Christiana D., 46, 55, 461, 486
Morgan, C. T., 240, 258
Morgan, W., 178
Morris, C. W., 54, 128-130, 141, 489, 539, 560
Morse, Nancy C., 485
Moses, 166-167, 175, 177, 179
Moses, L. E., 289, 312, 314-315, 321, 323, 334
Moshman, J., 304, 334
Mosier, C. I., 343, 348, 368
Mosteller, F., 65, 73, 77, 83, 87, 89, 289, 343, 345, 363, 368
Motz, Annabelle B., 227, 257
Mowrer, O. H., 59, 64-66, 71-73, 82, 84-86, 88-89, 114, 120, 131, 141, 145, 176, 501, 519, 550-552, 560-561
Muensterberger, W., 171-172, 178, 180
Muhlhan, Gertrude, 500, 519
Mukerjee, R., 49
Müller, G. E., 44, 55
Müller, J., 144
Münsterberg, H., 43, 47, 55
Murchison, C., 43, 49-52, 55, 138, 173, 179, 221, 366, 403, 561
Murdock, G. P., 62, 86, 89, 523, 525-526, 528-531
Murphy, G., 6, 48-50, 52, 55, 133, 141, 182, 223, 250, 257, 276, 288, 364, 368, 407

Murphy, Lois B., 21, 48-50, 55, 276, 288
Murray, H. A., 12, 46, 54-55, 70, 89, 145, 178, 237, 255, 286, 288, 461, 486
Mussen, P. H., 342, 368
Myers, C. S., 35, 336, 366

Naboisek, H., 229, 257
Nacht, S., 169, 179
Napoleon, 129
Neiman, L. J., 224, 257
Nelson, C. J., 512, 522
Nelson, M. J., 246
Newcomb, T. M., 49-51, 55, 173-174, 179-180, 182, 221, 223, 232, 244, 257, 276, 288, 342, 368, 409, 422, 431, 447, 461, 486, 531
Newstetter, W. I., 409, 422, 447
Newton, I., 181-182, 532
Neyman, J., 291, 334
Nietzsche, F., 14, 55
Nissen, H. W., 532, 534, 540, 542, 551, 560
Niven, J. R., 353, 365, 368
Noland, E. W., 355, 368
Norman, R., 199, 220
Northway, Mary L., 405, 412, 414, 422, 426-427, 434-435, 437, 447
Notcutt, B., 237, 257
Nottingham, Elizabeth K., 228, 257
Novikoff, A. B., 125, 141
Nowlis, Helen H., 81, 89
Nowlis, V., 81, 89, 486
Nunberg, H., 168, 179
Nuttin, J., 123, 141

Oersted, H. C., 181
Ogburn, W. F., 4, 55, 531
Ogden, C. K., 16, 55
Ojemann, R. H., 491, 521
Olds, J., 82, 85, 89
Olesen, Barbara A., 447
Olson, W. C., 371, 392-393, 404
Opler, M. E., 171, 179
Ort, R. S., 228, 257
Osburn, H., 385, 404
Osgood, C. E., 75, 84, 89, 128-129, 141
O.S.S. Assessment Staff, 342, 368, 451, 486
Ossorio, A. G., 389, 403-404
Ovsiankina, Maria (*See* Rickers-Ovsiankina, Maria)

Pace, C. R., 341, 368, 486
Paget, Violet (*See* V. Lee)
Palmer, Georgiana P., 498, 512, 521
Papania, N., 222
Pareto, V., 11, 17, 52, 54-55, 78
Park, O., 536, 559
Park, R. E., 45, 55, 224, 238, 257
Park, T., 536, 559
Parry, H. J., 480, 486
Parsons, Elsie C., 170, 179
Parsons, T., 5, 11, 38, 40, 55, 61, 69-70. 78, 89-90, 142, 172, 179, 182, 190, 194, 222, 223-225, 227, 253, 257
Parten, Mildred B., 456, 486
Paterson, D. G., 465, 485
Paul, B. D., 473, 486
Pavlov, I. P., 57, 61, 64, 72, 89, 130, 169
Payne, S. L., 341, 368, 456, 465, 468, 485-486
Pear, T. H., 21, 55
Pearson, K., 48, 291, 318, 334, 415-416, 437
Pelley, W. D., 502
Pepinsky, H. B., 409, 415-416, 435, 440, 446-447
Pepinsky, Pauline N., 406, 421, 423, 447
Pepitone, A. D., 212, 222
Perkins, D. T., 88
Perlmutter, H. V., 212, 220
Perry, A. D., 418, 446
Perry, R. B., 39, 55
Perry, S. E., 175, 180
Peterson, D. R., 233, 248, 256
Phadke, N. S., 49
Piaget, J., 38, 55, 110, 115, 132-137, 141, 242-243, 257, 451, 471-472, 486
Pilzecker, A., 44, 55
Pintner, R., 351, 368, 433, 447
Plath, O. E., 536-537, 561
Plato, 3, 5-6, 16, 18-19, 29, 32, 53, 55-56
Poe, E. A., 497
Poisson, S. D., 326-328
Polansky, N. A., 210, 220-222, 374, 387, 399-400, 404, 423, 447
Pomeroy, W. B., 480, 485
Pool, I. deS., 488-489, 508, 511, 517, 521
Posner, Rita, 435, 445
Postman, L., 57, 59, 87, 89, 110, 113-114, 117, 138, 141, 357, 368
Potter, Elmer H., 507-508, 519
Powell, R. M., 412, 428, 430, 446-447
Pratt, C. C., 92, 102, 117, 141
Prentice, W. C. H., 201, 222
Prince, M., 26
Proctor, C. H., 405, 411-412, 414-415, 419, 421, 428-429, 446-447
Proshansky, H. M., 342, 368, 486
Pumpian-Mindlin, E., 178-179
Punke, H. H., 511, 521

Quarrington, B., 412, 426, 447
Queener, E. L., 49, 89, 173, 179
Quetelet, A., 48

Radke, Marian J., 472, 480, 486
Rainoff, T. J., 491, 521
Ramsey, G. V., 138-139
Rapaport, D., 138-139, 168, 179, 184, 221
Rapaport, Dorothy, 222
Ratliff, J. A., 437, 446
Reckless, W. C., 49

Redl, F., 210, 222, 374, 387, 404, 423, 447
Reik, T., 172, 180
Reinhardt, J. M., 49
Reisman, D., 172, 180
Reiss, A., 514, 521
Révész, G., 539, 561
Reymert, M. L., 138, 141, 220
Ribot, T., 20-21, 55
Rice, S. A., 5, 52, 55, 477, 486
Richardson, Helen M., 432, 447
Richardson, M. W., 339
Richmond, Winifred, 434, 447
Rickers-Ovsiankina, Maria, 128, 139, 201, 222
Rickert, Edith, 490, 521
Riesen, A. H., 534, 551, 560-561
Riesman, D., 462, 486
Riker, B. L., 362, 368
Ritchie, B. F., 107, 141
Rivers, W. H. R., 35, 55, 170, 180, 235, 257
Roberts, J. M., 486, 530
Robinson, B., 533, 560
Robinson, J. H., 17, 55
Roby, T. B., 89
Roethlisberger, F. J., 438, 451, 486
Rogers, C. R., 18, 55, 115, 121-122, 141, 244, 257, 385, 404, 451, 465-466, 486
Rogers, Maria, 438, 447
Rogers, W. C., 515, 521
Roheim, G., 35, 55, 171-172, 176, 178-180, 520
Rohrer, J. H., 54, 140-141, 559
Rokeach, M., 247, 257
Romney, K., 486
Roosevelt, F. D., 496, 508, 522
Roper, E., 457
Rorschach, H., 178, 249-250, 326, 333, 382, 420, 435, 447, 452, 460, 499, 521
Rosander, A. C., 341, 368
Rose, A. M., 227, 257
Rosen, E., 95, 140, 183, 222
Rosen, S., 210, 221
Rosenbaum, M., 256
Rosenberg, B. G., 230, 244, 258
Rosenstock, J., 461, 486
Rosenzweig, S., 201, 222, 236, 252, 257, 342, 445, 550, 561
Roslow, S., 364, 368
Ross, E. A., 3-4, 9, 19, 27, 30-31, 48, 55
Ross, Helen, 179
Ross, I. C., 418, 447
Ross, R. T., 88
Ross, S., 62, 90, 534, 561
Rossi, P. H., 361, 368, 464
Rothmann, E., 128, 141
Rousseau, J. J., 15, 56
Royce, J., 36, 55
Rubin, E., 98
Ruesch, J., 138
Rundquist, E. A., 365
Runion, H. L., 498, 512, 521
Ruskin, J., 11
Russell, B., 6, 14, 55

Sabine, G., 6, 55
Saffir, M. A., 260, 288, 347-348, 364, 368
Sait, E. M., 208, 220
Salter, Patricia, 494, 509-511, 515, 517, 519
Sanders, I. T., 441, 447
Sanderson, D., 405, 447
Sanford, F. H., 342, 368, 461, 486
Sanford, R. N., 176, 485
Sapir, E., 40, 170-172, 180, 256
Sarason, S., 236, 257
Sarbin, Genevieve, 223
Sarbin, T. R., 223-224, 227, 229-230, 232, 234, 237-238, 240, 244, 246-249, 253, 255, 257-258
Sargent, J., 409, 437, 446-447
Sargent, S. S., 49, 224, 258, 550, 561
Schachter, S., 215-217, 220, 222, 418, 430, 445, 485
Shaeffle, A., 35, 55
Schanck, R. L., 49
Schauer, G., 440, 447
Scheerer, M., 91, 126, 128, 131-132, 139, 141
Scheffé, H., 304, 334
Scheidlinger, S., 167, 180
Scheler, M., 20-21, 55
Schilder, P., 111, 115, 133, 141, 240, 258
Schlesinger, H., 113, 140
Schlipp, P. A., 51
Schlosberg, H., 64-65, 71, 80, 89-90, 257, 534, 560
Schmidt, J. K. (*See* M. Stirner)
Schmidt, K. P., 536, 559
Schneider, B. H., 117, 141
Schneider, D. M., 178
Schneider, L., 175, 180
Schneirla, T. C., 124-125, 141, 532, 534-536, 538-539, 541, 560-561
Schopenhauer, A., 16, 55
Schramm, W., 521
Schuessler, K. F., 248, 258
Schultz, J. H., 237, 258
Schutz, W., 515, 521
Schwartz, M., 351, 367
Schwartz, S., 256
Scott, J. P., 541, 561
Seagoe, Mary V., 433, 447
Sears, Pauline S., 80, 89, 208, 221, 486
Sears, R. R., 69-71, 81-84, 86-89, 145, 170, 176, 180, 243, 248, 258, 286, 288, 481, 486, 524-525, 531, 540, 543, 561
Seeman, M., 416, 447
Seligman, C. G., 170, 180
Sellars, R. W., 141
Semmes, Josephine, 534, 560
Sengupta, N. N., 49
Seward, J. P., 76, 89, 128-130, 141
Shakespeare, W., 497-498, 521

Shanas, Ethel, 491, 521
Shand, A., 46, 55
Shapiro, J., 477-478, 486
Shartle, C. L., 409, 439, 448
Shaw, G. B., 28
Sheats, P., 232, 256
Sheatsley, P. B., 465, 471, 486
Sheffield, F. D., 76, 88-89, 277, 287, 450, 485
Sherif, M., 49-50, 54, 109-110, 140-141, 342, 351, 367-368, 559
Shils, E. A., 5, 40, 55, 61, 69, 89-90, 142, 172, 179-180, 190, 194, 222, 223-225, 227, 253, 257
Shipley, T. E., 85, 89
Shirley, Mary M., 241, 258
Shneideman, E. S., 499, 521
Shoobs, N. E., 437, 447
Shor, J., 506, 520
Shvernik, N. M., 499
Sickles, W. R., 240, 258
Sidis, B., 25-26, 30, 55
Siegel, L., 409, 435, 447
Sighele, S., 25, 30, 55
Silva, A. L. M., 237, 257
Simmel, E., 169, 174, 180
Simmel, G., 30, 55
Simmel, Marianne, 105, 139, 210, 212, 220
Simmons, L. W., 523, 531
Simon, C. W., 287
Simon, H. A., 182
Simpson, G. G., 125, 141
Sisson, E. D., 362, 368
Skinner, B. F., 63-66, 71, 73-74, 85, 90, 92, 96, 102, 141, 286, 288
Sletto, R., 365
Sliosberg, S., 202, 222
Sloane, E. H., 93, 141
Smelser, N., 431, 447
Smith, A., 10, 19-21, 55
Smith, B. L., 520
Smith, G. H., 342, 368
Smith, H. L., 478, 486
Smith, J. J., 49
Smith, Kate, 507-508
Smith, M., 431, 447
Smith, M. B., 238, 258
Smith, R. G., Jr., 364, 368
Smith, S., 76, 90
Smith, W. I., 62, 90, 534, 561
Smucker, O., 414-415, 447
Smythe, D., 493, 521
Snedecor, G. W., 268, 285, 288, 326, 334, 345, 368
Snyder, W. U., 385-386, 404
Snygg, D., 101-102, 122, 141
Socrates, 29
Solomon, R. L., 60, 72-74, 79-80, 85, 88-90, 257, 276-277, 288, 420, 422, 425, 427, 431-432, 445-446, 534, 560
Sorel, G., 17, 55
Sorokin, P. A., 5, 18, 21, 51, 55
Spearman, C. E., 232-233, 339, 392
Speier, H., 174, 179, 508, 520
Spence, K. W., 57, 59, 61-62, 69-70, 74, 78, 80, 87, 90, 91-93, 96, 124, 130-131, 141, 184, 222, 336, 366
Spencer, H., 4-5, 11, 19-20, 32, 43, 48, 56
Spenser, E., 497
Speroff, B., 434, 438, 448
Sperry, Margaret, 82-83, 88
Spier, L., 256
Sprott, W. J. H., 49
Sprowls, J. W., 5, 56
Spurgeon, Carolina F. E., 490, 497, 521
Stalin, J., 499, 521
Stanley, J. C., Jr., 58, 88
Stanton, A. H., 175, 180
Stanton, F. N., 519, 521
Starbuck, E. D., 48, 56
Steele, Mary, 408, 428-429, 446
Steinthal, H., 3, 35-36, 54
Steinzor, B., 380-381, 396, 404
Stellar, E., 240, 257-258, 534, 560
Stember, H., 469, 471, 486-487
Stephenson, R. W., 247, 258
Stern, W., 93-94, 106, 111, 132, 134, 141-142, 371
Stevens, N. E., 491, 521
Stevens, S. S., 61, 87-90, 92, 99, 140-142, 335, 337, 368, 560
Stevenson, R. L., 26
Stewart, F. A., 408, 429-430, 441, 448
Stewart, Isabel A., 408, 448
Stewart, Jeannie C., 541, 561
Stewart, M., 516-517, 521
Stewart, Naomi, 471, 476-478, 487
Stine, Margaret E., 174, 179, 435, 447
Stirner, M., 14, 56
Stogdill, R. M., 409, 425, 439, 448
Stone, C. P., 87, 89, 287, 560
Stouffer, S. A., 18, 56, 253, 258, 276-277, 288, 329-331, 334-335, 343, 356, 367-369, 450, 465, 476, 482, 487
Stratton, G. M., 56
Strauss, A. L., 49, 51, 87-88, 126, 140, 257
Strodtbeck, F. L., 486
Stroud, J. B., 62, 90
Suchman, E. A., 335, 355, 357-358, 369
Sullivan, H. S., 173, 182
Sullivan, P. L., 228-229, 258
Sumner, W. G., 4
Sussman, Leila A., 494, 517, 521
Sutherland, J. D., 439, 448
Suttie, I. D., 18, 56
Sutton, F. X., 31, 51
Swanson, G. E., 174, 179-180, 531
Swed, Frieda S., 289, 322, 334
Symonds, P. M., 238, 258, 468, 487
Szekely, L., 109, 142

Taeuber, Irene B., 490, 511, 515, 517, 522
Taft, R. A., 494, 519
Tagiuri, R., 409, 426-427, 429, 435, 445, 448
Talbot, E., 409, 438, 448
Tannenbaum, R., 409, 438, 446, 448
Tarde, G., 3, 8, 22, 24, 30, 48, 56, 81
Taylor, D. W., 287
Taylor, W. S., 145, 180
Terman, L. M., 331, 334
Thelen, H., 395-396, 404
Thibaut, J. W., 215-217, 220, 222, 398, 404, 485
Thistlethwaite, D., 118, 142, 342, 369
Thomas, Dorothy S., 371, 398, 404
Thomas, W. I., 43-45, 56
Thompson, W. R., 532-534, 551, 561
Thorndike, E. L., 13, 41, 56-57, 61, 63, 76, 90, 114, 142, 496
Thorndike, R. L., 279-282, 288
Thorpe, W. H., 537, 561
Thouless, R. H., 49
Thrasher, F. M., 437, 448
Thurnwald, R., 35, 56
Thurstone, L. L., 48, 56, 78, 260, 264, 287-288, 335, 341, 344-345, 347-348, 350-351, 360, 362, 364-369, 400, 409
Tinbergen, N., 63, 90, 535, 561
Tippett, L. H. C., 263, 285, 288
Titchener, E. B., 10, 12, 26, 44, 56, 144
Toby, J., 253, 258
Tocher, K. D., 312, 334
Tolman, E. C., 61, 72-73, 78, 90, 92-94, 96, 104, 108, 110-111, 113, 117-118, 122, 124, 129, 142, 145, 169, 172-173, 182-183, 188, 190, 207, 222, 226, 245, 258, 286, 288, 538, 541, 561
Tomkins, S. S., 145, 180
Torgerson, W. S., 335, 337, 369
Travers, R. M. W., 362, 369
Travis, L. E., 47, 56
Triplett, N., 46-47, 56
Troland, L. T., 12, 56
Trotter, W., 20, 41, 56
Tryon, R. C., 533, 561
Tucker, L. R., 352, 369
Tukey, J. W., 289, 304, 326-327, 333-334
Turner, F. J., 493
Tylor, E. B., 523, 528, 531

Urdan, J. A., 434, 446

Vaihinger, H., 236, 258
Valentine, C. W., 551, 561
Vallance, T. R., 85, 90
Vanatta, E. L., 409, 435, 447
Van Dyne, Virginia E., 433, 448
Van Spanckeren, W. J., Jr., 244, 258
Vaughan, C. E., 6, 56
Vaughan, M. F., 49
Veblen, T., 175, 180
Veitch, J., 53
Vernon, P. E., 76, 87, 237, 432
Veroff, J. A., 85, 89
Vicary, J. M., 461, 487
Vico, G., 3, 56
Vigotsky, L. S., 126, 142
Vogt, E. Z., 530
von Bertalanffy, L., 95, 111, 142
von Frisch, K., 537-538, 561
von Holst, E., 111
von Humbolt, K. W., 35
von Senden, M., 100, 142, 534, 553, 561
Voroshilov, K. E., 499
Vreeland, F. M., 430, 448

Wade, M., 130, 140
Waelder, R., 28, 56, 168, 174, 180
Wald, A., 290, 321
Walk, R. D., 85, 90
Walker, Helen M., 263, 269, 288
Wallace, D. L., 289, 304, 320, 334
Wallach, H., 92, 99, 142
Wallas, G., 18, 41, 56
Waller, W. W., 172, 180, 223, 249, 258
Wallin, P., 223, 228, 258
Wallis, W. A., 289, 320, 328, 330-331, 334
Wallis, W. D., 39, 56
Walworth, A., 491-492, 517, 522
Waples, D., 488-489, 496, 508, 511-512, 519-520, 522
Wapner, S., 105, 117, 131-132, 142
Ward, C. H., 17, 56
Warden, C. J., 125, 142
Warner, W. L., 228, 235, 258
Warren, H. C., 27, 45, 56
Warren, R. L., 252, 258
Washburn, Margaret F., 44, 56
Wasmann, E., 536, 561
Waterhouse, I. K., 87
Watson, G., 47, 56, 219
Watson, J. B., 42, 56, 57-58, 73, 90, 92, 94, 550
Watson, P. O., 60, 85, 88
Wayne, I., 492, 504, 511, 521
Weaver, C. L., 289, 291, 333
Weiss, P., 132, 142
Weiss, W., 84, 88, 90
Welles, O., 30
Wells, F. L., 173
Werner, H., 105-106, 117, 124, 126, 128, 132-133, 136-137, 142, 239-240, 250, 258
Wertheimer, Max, 98, 108-109, 134, 142
Wertheimer, Michael, 100, 142
Weschler, I. R., 342, 369, 409, 412, 438, 446, 448, 461, 487
Wever, E. G., 109, 142
Whaley, F., 222
Wheeler, G. C., 523, 531
Wheeler, W. M., 534, 536-537, 561
Wherry, R. J., 352, 369, 425, 439, 448

White, C. T., 65, 90
White, L. A., 38, 56, 126, 142, 525, 531
White, R. K., 48, 54, 207, 209, 214-215, 218, 221-222, 496, 501, 510, 522
Whiteman, P. H., 534, 559
Whiting, Beatrice B., 523, 528, 530-531
Whiting, J. W. M., 58, 71, 83-84, 86-87, 90, 486, 523-526, 528-531, 551, 561
Whitney, D. R., 290, 312, 315-318, 323, 325, 334
Whittemore, I. C., 47, 56
Whitton, J. B., 520
Wiehe, F., 209, 222
Wiener, N., 34, 56, 76
Wigdor, Blossom T., 435, 447
Wilbur, G., 171-172, 178, 180
Wilcoxon, F., 289-290, 314-315, 334
Wild, J., 56
Wilkinson, N. M., 393, 404
Wilks, S. S., 308, 334, 352-353, 369
Willey, M. M., 489, 494, 511, 515, 522
Williams, D. C., 532
Williams, J. D., 230, 258
Williams, J. M., 49
Williams, S. B., 409, 425, 439, 448
Williams, W., 14, 56
Willkie, W. L., 496
Willoughby, R. R., 173
Wilson, T., 289
Wilson, W., 498, 521
Winch, R. F., 172, 180
Windelband, W., 6, 56
Winslow, C. N., 432, 448
Wise, G. W., 524-525, 531
Wishner, J., 106, 141
Wissler, C., 24, 56
Withall, J., 395, 404
Witkin, H. A., 109, 117, 142
Wittreich, W. J., 113, 142
Wolfenstein, Martha, 504-505, 518, 522
Wolff, K. H., 55
Wolfowitz, J., 290, 321
Wolman, S., 441, 448
Wood, L., 103, 142
Woodward, J. L., 469, 487, 505, 508, 522
Woodworth, D. G., 223
Woodworth, R. S., 41, 56, 74, 116, 118, 142, 145, 261, 288, 551, 561
Wordsworth, W., 497
Wormhoudt, A., 500, 522
Wright, Beatrice A., 91, 198, 219
Wright, G. O., 83, 90, 523, 528, 531
Wright, H. F., 107, 142, 182-183, 193-194, 219, 221-222, 286, 288, 388-389, 403-404, 484-485
Wright, Q., 512, 522
Wright, R., 501
Wrighter, Joyce, 429, 431, 433, 445
Wulff, J. J., 89
Wundt, W., 5, 8, 10, 27, 35-37, 44, 46, 56, 144, 370
Wyant, Rowena, 509, 522
Wyatt, D. F., 478, 487
Wyatt, T., 497
Wynne, L. C., 72-73, 85, 90

Yakobson, S., 491, 522
Yates, F., 262, 273, 287, 314, 333
Yeats, W. B., 497
Yerkes, A. W., 554, 561
Yerkes, R. M., 104, 130, 539, 543-545, 549-551, 554, 561
York, R. H., 440, 445
Yoshioka, J. G., 554, 560
Young, K., 45, 49, 56, 224, 238, 243, 258
Young, P. T., 115, 142
Youtz, Cleo, 289
Yule, G. U., 498, 522

Zander, A., 394, 404
Zeigarnik, Bluma, 112, 200-201, 218, 222
Zelen, S. L., 244, 256
Zeleny, L. D., 409, 414, 422, 426-428, 437, 440, 448
Zeller, A. F., 201, 222
Zener, K. E., 109, 142, 257
Zilboorg, G., 178
Zimmerman, C., 357, 368
Znaniecki, F., 43-45, 48, 56
Zubek, J. K., 532
Zuckerman, S., 541, 561

SUBJECT INDEX

Abstraction, 127-128, 132, 136-137
Accident proneness, 434, 438
Achievement, and sociometric status, 431
Acquisition, theories of, 63-65
  See also: learning, conditioning, stimulus-response theory
Acting, 233, 249
Action, 224-228, 230
Action research, 212-213
Adaptation, 133-134
Additive property, 336-337
Adjustment, 227
  social, sociometric measures of, 426-427, 435
Adolescence, 198
Afferent interaction, 100
After-only design, 277
Age, and sociometric status, 430
Aggression, 71, 83, 159-160, 168-169
  in chimpanzees, 546-550
Altruism, 541-543
Ambivalence, 166, 206
Analogical theory, 18-19, 32-34
Analysis of variance, 267-268, 290, 304-307, 320-321, 327-328
  error variance, 327-328
  one-way classification, 304-305
  ranked one-way, 290, 320-321
  short-cut allowances in, 290, 304-307
  two-way classification, 305-307
Anger, 549-550
Animal magnetism, 25
Animal psychology, 124-125, 532-558
  relevance of, 532-534
  studies of intelligence in, 533-534
Animation, 106
Annoyance, 245
Antecedent-to-response analysis, 66
Anthropology, 86
  and psychoanalysis, 170-171
  use of interviews in, 473-475
  See also: cross-cultural method
Anticathexis, 147, 156, 163
Anticipatory phenomena, 62-63, 116
Antisocial behavior, 248
Anxiety, 85, 160-163, 530
  See also: psychoanalytic theory, anxiety
Anxiety reduction, 65
Aphasia, 128
Apparent behavior, 104-107, 212
Applied research, 260
Approach, constructive *vs.* classificatory, 187-188
Arc sin transformation, 345
Aristotelian thinking, 185
Ascendance, 391-392
"As if" behavior, 234, 236-238, 249
Assignment therapy, 437-438
Association, 37, 67, 73-74, 152, 199-200
  by similarity, 37
Associationism, and suggestion, 26
Associative chain theory, 127
Attitude, 43-46, 97, 119-120, 238, 335-336, 357-358, 431-432, 476-478, 534, 543-548
  definition of, 43, 45, 335-336
  effect of interviewer's, 477-478
  intensity of, 357-358
  recognition of, 544-545
Attitude change, 84, 216-217
Attitude measurement, 48, 335-366
Attitude scale, zero point of, 357-358
  See also: attitude measurement, scaling
Attitude universe, 336, 340
Aufgabe, 43, 45-46, 110, 128
Aussage test, 342
Authoritarianism, 209, 247
Autism, 135, 153

Babbling, 126
Balanced Latin square design, 285
Barrier, 198, 204, 206
Before-after design, 277
Behavior, biosocial *vs.* psychosocial, 535
  in field theory, 191-192
Behavior disposition, 129
Behavior setting, 194, 219
Behavioral contagion, 210, 387-388
  See also: imitation
Behaviorism, 57, 91-93, 129-130
  See also: conditioning, stimulus-response theory
Bias, 351
Binomial distribution, 291-292, 326
Body image, 111-114, 240
Boundary, 196
Brain injury, 128
Bureaucracy, 249

Capacity, man *vs.* animal, 125
Capacity fulfillment, 115-116
Caste, insect, 535-536
Castration fear, 161, 166
Category, discrete *vs.* continuous, 374-375
Category sets for observation systems, 372-375
Category system, Bales', 376-378
  conference research problem solving, 378-380
  Steinzor's social interaction, 380-381
Cathexis, 146-147, 156, 163

Causation, 103-104, 186-187, 194
  systematic *vs.* historical, 186-187
Change, socially induced, 209-210, 217
Character formation, 165
Chi-square test, 293-294, 329, 345
Child development, 83-84
  See also: Cognitive development, development
Child psychology, 391-393, 471-473
Child-rearing practices, 239, 243, 524-525
Chimpanzees, social behavior of, 540, 542-550
Choice, 207
Circular reflex, 126
Co-acting group, 47
  See also: face-to-face group, small group theory
Codification, 99
Coefficient of homogeneity, 357
  of reproducibility, 355-356
Coercion, 194-195
Cognition, and conation, 114-123
  definition of, 94
  and manifest organization, 103-104
  and reinforcement, 117-118
  role in behavior of, 91-93, 107-108
  and secondary reinforcement, 118-120
Cognitive control, 117
Cognitive development, 132-137, 238-244
Cognitive organization, 18, 23, 28-29, 108-109
  and social perception, 104-107
Cognitive structure, 197-198, 207, 211
Cognitive theory, 91-137
  vs. stimulus response theory, 93, 96-97, 113-115, 118-120
Cohesiveness, 214-216
Collective, consciousness, 35
  representation, 37-38
  unconscious, 34
Common enemy, 14
Common segment theory, 39-40
Commonality, 101, 194-195
Communication, 125-127, 216-217, 537-540
  in insect societies, 537
  See also: content analysis, information theory, language
Competition, 202
Conation, *see:* motivation
Concept, 125-130, 539
  in field theory, 187
Concept formation, 126-127, 130-132, 242
Concreteness, 127-128, 136-137
Conditioned response, type R *vs.* type S, 64
Conditioning, 23, 26, 59-61, 96, 118-119, 130-131
  higher order, 62
  and imitation, 23
  spontaneous recovery in, 61
  and suggestion, 26
  temporal relations in, 61
  See also: association, learning, stimulus-response theory
Conditions, experimental, arrangement of, 270
Confidence limits, 305
Conflict, 197-198, 205-207, 248, 251-253
Conformity, 21-24, 216-217
Confounding, in experiments, 269, 276
Conscience, 153-154
Consciousness (conscious), 57, 148, 155-156
  of kind, 20
  See also: Psychoanalytic theory, conscious
Consistency, 121-122
Consonance, 198
Constancy, 98-99, 239, 251-252
Contemporaneity, 186
Content analysis, 488-518
  audience of content, 503-505
  categories of, 510-512
    formal categories, 511-512
    substantive categories, 510-511
  conditions for exact counting, 512-514
  definition of, 488-489
  effects of content, 505-507
    focus of attention, 505-506
    response to communication, 506-507
  formal characteristics of content, 495-498
    exposure of propaganda, 495-496
    readability, 496-497
    stylistic features, 497-498
  history of, 489-490
  inference in, 516-518
  and personality measurement, 500-501
  producers of content, 498-503
    detection of propaganda, 501-502
    determination of psychological states, 499-501
    intent of communicator, 498-499
    political and military intelligence, 502-503
  and projective tests, 499
  qualitative analysis, 512
  recording unit *vs.* context unit, 507
  reliability of, 514-515
  sampling in, 515-516
  substantive characteristics of, 490-495
    communication standards, 494
    comparison of communication media, 492-493
    content *vs.* its objectives, 493-494
    development of scholarship, 491
    international differences in content, 491-492
    as a research tool, 494-495
    trends in content, 490-491
  technical problems of, 512-518
  units of, 507-510
    character, 509
    interrelationship of units, 509
    item, 509
    space and time units, 509
    theme, 508-509
    unit of classification *vs.* unit of analysis, 507-508
    word, 508

  uses of, 490-507
  validity of, 504-505
Contiguity theory, 63, 73-77
Continuity in learning, 75
Control group, function of, 276-277
  number of subjects in, 277
Control group design, extension of, 276-277
Conversation of gestures, 22
Cooperation, 21, 202, 540-541
Copying, 23
  See also: imitation
Correction for continuity, 309
Covariance techniques, 274
Craze, 31
Critical ratio test, 302-303
Critical research, 260
Cross-cultural method, 523-531
  and behavior science, 523-525
  history of, 523
  research in, 528-531
  sampling in, 526-527
  units of, 527-528
  See also: anthropology
Crowd psychology, 25-26, 28-31
Cue-drive learning, 72
Cultural determinism, 38-39
Cultural relativism, 101
Culture, 36-37, 85-86, 101, 113, 167-168, 208, 224, 239, 266, 430-431
  effect of on cognition, 101, 113
  effect of on level of aspiration, 208
  and personality, 239
  and psychological processes, 266
  and sociometric status, 430-431
Culture trait, 527
Culturology, 38, 525-526
Curiosity, 118
Custom, 526-528
Cybernetics, 34

Death instinct, 146, 149, 156-159, 169
Death wish, 235
Decalogue, 14-15
Decentering, 133-134
Decision making, 207, 217-218
Defense mechanism, 162-164
  See also: ego defense
Delinquency, 437
Democracy, and crowd behavior, 30
Demographic variables, and sociometric variables, 429-430
Derivations, doctrine of, 17
Descriptive analogy, as theory generalization, 79-80
Determination, proximal *vs.* distal, 95-97
  relational, 98-99, 130
Determining tendency, 27, 110, 199-200
Development, psychological, 203-204
  See also: child development, cognitive development, differentiation, maturation
Developmental stage, 133-137
Deviation, 217
Differentiation, 132, 192, 196, 203-204, 239-240
  of the life space, 192
  of regions, 196
Diffusion, 24
Dimension dominance, 218
Direction, in field theory, 189, 197, 199
Directionality of result, 330
Discontinuity, in learning, 75
Discreteness, 136
Discrimination, 62, 130-131
Displacement, 82-83
Disposition, behavior, 129
Dissociation of consciousness, 25-26, 29
Distal determination, 95-97
Distance, in field theory, 197
Distinguished path, 189, 196, 199
Dominance, 392, 541
Drive, 60, 68, 70-71, 114-115
  effect of increase in, 68
  learnable, 70-71
  viscerogenic, 114-115
Drive reduction, 63-64, 72, 116
Drive-response learning, 72
Drive strength, 67-68
Dual theories of acquisition, 64-65, 71-73
Dyadic interaction, 30, 83, 86-87, 384-387, 540-541
Dynamic approach, in field theory, 188-189
Dynamogenesis, 25, 46

Echo principle, 23
Ecology, psychological, 185, 193-195, 388-389
Economic man, 10-11, 553
Ecstasy, 235
Education, and sociometry, 430, 436-438
Ego, 111-114, 117, 120-123, 148, 150-155, 159, 162, 168-169, 238, 254
  Freudian, 148, 150-155, 159, 162, 168-169
  vs. self, 238
  See also: body image, psychoanalytic theory, ego, self
Egocentrism, 135
Ego defense, 14, 162-164, 251-252
Ego ideal, 153
Ego instinct, 157
Ego involvement, 201
Egoism, 13-15
Electromagnetism, 181-182
Emergence, 37, 125-126
Emotion, 65-67, 122-123, 553-558
  and learning, 65-67
  See also: feeling, mental qualities, specific emotions (i.e., anger)
Emotional learning, 71
  speech, 126

Empathy, 20, 23, 231-232, 237-238, 246-248, 409, 427, 539-540
Energy, 145-147, 153, 156
Enlightenment, The, 15-16
Environment, 189-190, 204-212, 533-534
  in field theory, 189-190, 204-212
Epiphenomenalism, 37
Equal-appearing intervals, method of, 347
Equalitarianism, 247
Equivalence of meaning, 452-453
Erogenous zone, 158
Eros, *see* life instinct
Eroticism, 158
Error, choice technique, 342, 461
  in rating, 271-272
  Type I *vs.* Type II, 262-263, 291
  variance, 278
Escape behavior, 206
Estimates of time, 409-410
Ethics, utilitarian, 10
Ethnocentrism, 247, 339
Event system, 40, 85
Evolution, 11, 19-20
Ex post facto design, 281-282
Excitation, 68
Excitement, 551-552
Experimental design, types of, 277-285
Experimental error, 268
  reduction of, 272-274
Experimental method, 50, 260-261
  beginnings of, 46-48
Experimental technique, 274-275
Experiments, planning and execution of, 259-287
Extinction, 63-65

F-test, 268, 272-273, 294-302
Face-to-face group, 47, 376-384
  See also: small group theory
Face validity, 340, 398, 422-423
Faces test, 341-342
Factor analysis, 352
Factorial design, 282-284
Fallacy of unitary explanation, 9
Family structure, 165-167, 429
Fear, 70-72, 548-552
  as a learnable drive, 70-72
Feeling, 155, 239
  See also: emotion
Field, psychological, 185
  See also: life space
Field study, 213
Field theory, 181-219
  basic concepts of, 189-207
  changes in psychological environment, 207-212
  definition of, 182
  dynamic concepts in, 199-207
  mathematics in, 189, 195
  metatheory of, 182-183
  in physics, 181-182
  psychological approach in, 183-185
  structural concepts of, 195-199
Fixation, 240
Flexibility, 136, 247, 250
Fluidity, 201
Folk mind, 35-36
  psychology, 35-37
Force, psychological, 188-189, 204-207, 209
Forced choice technique, 362-363
Fractional anticipatory goal response, 62
Frame of reference, 109-111, 194
Friendship, 541-543
Frustration, 83, 147, 203-204, 549-551
  and learning, 66-68
Functional autonomy, 120
  See also: motives, transformation of

Generalization, 82-83
  of results, 265, 274
  of theory, 77-81
  See also: stimulus generalization
Generalized other, 377-378
Genetic reductionism, 124-132
Geographic environment, 97-98
Genotype, 187
Gestalt psychology, 18, 23, 28-29, 40, 97-99, 103-113, 119, 122-123, 130-132
Gesture, 242
Gradient of generalization, 61-62
Gravitation, 181-182
Gregariousness, 19-20
Group atmosphere, 48, 209, 380-381
Group decision, 217-218
Group, definition of, 214
Group dynamics, 212-219
Group fallacy, 39
Group mind, 18-19, 31-40, 213-214
  analogical theory of, 18-19, 32-34
  as collective representations, 37-38
  as collective unconscious, 34
  common segment theory of, 39-40
  criteria of, 32, 39
  as cultural determinism, 38-39
  as folk psychology, 35-37
  as objective mind, 34-35
Group participation scale, 409
Group preference record, 409
Group psychology, 167
  See also: small group theory
Group structure, 218, 428-429
  sociometric measures of, 428-429
Guess who technique, 410
Guilt, 82, 154, 161-162
Guttman type scale, 353-358, 361-362

H statistic, 320-321
H technique, for scaling, 356

Habit, 17-18, 42-43, 67, 96, 113, 238, 526-528
  family hierarchy, 96
  strength, 67
Hallucination, 149
Halo effect, 362-363
Hedonism, 10-13, 114, 120
Hedonistic calculus, 10, 16
Heredity, 533-534
Heteronomous event, 194, 210
Heteroscedasticity, 325
Hierarchy, of function, 131-133, 136
  of needs, 116
  of roles, 253
Hodological space, 189, 196-197, 199
Homeostasis, 114-115, 238-239
Homogeneity, of a scale, 339-340
  of variance test, 275
Hostility, 548-552
Housing, 441
Human nature, 552-558
Human Relations Area Files, 529-530
Hunger, 234
Hypnosis, 25-27, 233-234, 236-237, 249-250
Hypochondriasis, 247
Hypothesis, 110, 113
  testing, 261, 332
Hypothetico-deductive method, 65
Hysteria, 234-235

"I" concept, 242-243
  See also: ego, self
Id, 148-150, 153-154, 163
Idea psychology, 23
Identical element, 37, 75, 129
Identification, 23-24, 28, 84, 121, 146-147, 153, 165-167, 227, 236
Identity of perception, 150
Ideology, 11, 13, 17
Ideo-motor theory, 23, 25
Image, in crowd behavior, 29-30
Imagery, 237
Imagination, 226-227, 236
Imitation, 18-19, 21-24, 81-82, 119, 210, 228, 387-388
  definition of, 24
  forms of, 22
  laws of, 22
Impulse, 149
Impunitiveness, 236-237
Index of social distance, 411-412
Individual differences, 70, 269
  postulate, 70
Individualism, 11
Industry, and sociometric techniques, 438
Inferiority complex, 14
Information theory, 34, 91, 99
Inhibition, 64, 68
Insects, social phenomena among, 534-537
Insight, 104, 108-109, 131, 247
Instinct, 18, 24, 41-42, 146, 148-149, 156-160, 169
  Freudian, 146, 148-149, 156-160, 169
  of submission, 24
  of thought, 18
  See also: Psychoanalytic theory, instincts
Institution, social, 165-167
  and group mind, 31
Instrumental conditioning, and imitation, 23
Instrumental response, 60
Integration, 392
Intelligence, 431, 533-534, 553-555
  and emotion, 553-555
  and sociometric status, 431
Intention, 116-117, 120-121
Interaction chronograph, 394-395
Interaction effects, 276-277, 282, 284
Interaction recorder, 394
Interaction theory, 223, 238-239, 376-378, 380-381, 384-385, 535
Interests, 120-121, 205
Interjection, 126
Internal validity, *see:* face validity
Interpersonal mechanisms, 389-391
Interrupted tasks, 200-202, 218
Intervening variable, 69, 78, 107, 335
Interview, 449-484
  as a research tool, 450-451
  content analysis of, 495
  definition of, 449
  funnel, 459
  nondirective probing in, 466, 478
  open *vs.* closed questions in, 457-459
  vs. other techniques, 482-484
  phrasing of questions in, 455-459, 476-477
  reliability of, 453-454
  role relationships in, 462-464
  standardized *vs.* unstandardized, 451-455
  validity of, 454, 479-482
Interview data, recording of, 467-470, 478-479
  coded *vs.* verbatim, 469-470
  field *vs.* office coding, 469
  field *vs.* office rating, 468-469
  memory *vs.* mechanical, 468, 470
Interviewer bias, 462, 475
  characteristics of, 471, 476-478
  error, sources of, 475-479
Interviewing, of children, 471-473
  creating rapport in, 464-465
  cross-cultural, 473-475
  direct *vs.* indirect, 460-462
  field administration of, 471
  historical roots of, 449-450
  techniques of, 465-467
Introjection, 121
  See also: identification
Invariance, 99
Inverse sine transformation, 326-328
Involvement, in a role, 233-235

Irrationalism, 15-18
Irreality, level of, 192-193, 236
Isomorphism, 98
Item, 338, 341-350, 363-364
  form of, 343-344
  monotone *vs.* nonmonotone, 343-344
  scaling of, 344-350
  selection of, 350, 364
  types of, 341-343
Item analysis, 351-352

Judgment, 153
Justice, 19, 168

Lability, 136, 250
Lady tea taster, 270-271
Laissez-faire, 10-11
La morale, 7-9
Language, 125-130, 242-243, 538-540
  See also: communication
Language differences, in interviewing, 473-474
Latency period, 158
Latent, class analysis, 360-361
  distance scale, 361-362
  learning, 117-118
  structure analysis, 352, 359-362
  variable, 335
Latin square design, 284-285
  balanced, 285
Laughter, 20
Law, of affective evolution, 15
  of effect, 11-13, 63, 114, 120-121
  of three stages, 4, 7
Leadership, 28-30, 209, 232-233, 382, 410-411, 424-425, 439, 541
  sociometric measures of, 424-425, 439
Learnable drive, 70-71
Learning, 59-77, 108-109, 117-118, 121, 123-125, 207, 242
  definition of, 67
  of emotions, 123
  vs. performance, 69
  processes of, 61-63
  See also: acquisition, cognitive organization, conditioning, extinction, stimulus-response theory
Level of aspiration, 208-209
Leviathan, 13
Libido, 146, 157, 169
Life instinct, 146, 157-159
Life space, 184-185, 189-195
  dimensions of, 192
Likert type scale, 351-353
Locomotion, direction of, 189
  in field theory, 196, 204
"Looking glass" principle, 238, 243
Love, 158

Manifest organization, 103-104
Mann-Whitney test, 290, 315-317
Marital relationship, 227
Matched-dependent behavior, 82
  See also: imitation
Matched groups design, 278-281
  regression effects in, 279-281
Matching problem, 290, 307-312
Maturation, 239
Meaning, 84-85, 126-127, 452-453
Measurement, 272, 336-337
  random errors of, 272
  See also: experiments, quantitative techniques
Mechanism, 188
Median test, 290, 314
Mediating reaction, 129
Membership motive, 215
Memory, 111-112, 152
Mental hygiene, 534
Mental qualities, 155-156
Mesmerism, 25
Metaphysical absolute, 14
Milieu, nonpsychological, 193-195
Military applications of sociometric techniques, 439-440
Model, 332-333
Molar *vs.* molecular unit, 63, 74, 94-109, 183
Molar unit, criticism of, 95
  varieties of, 94-95
Monoideism, 25, 27
Moral anxiety, 161
Moral realism, 38
Morale, sociometric measures of, 428, 440
Motivation, 114-123, 156, 188, 374, 380-381, 397, 543-548, 550-552
  inferred from behavior, 374, 380-381, 397
  intrinsic, 122, 551-552
Motives, transformation of, 15, 70-71
Multivariate analysis, 332
Mutually shared field, 106-107
Mystic writing pad, 152

Narcissism, 146-147
National character, 492
Need, 114-117, 156-157
  cognitive, 114-116
  effect of on perception, 117
  primary, 114-115
  See also: drive
Neurosis, 252
Neurotic anxiety, 160-161
"New look" experiments, 113, 117
Nondirective counseling, 18
Nonparametric methods, 290-291, 312-323
Nonpsychological milieu, 193-195
Normal test, 302-303
Null hypothesis, 291, 293

Objective anxiety, 160
Objective mind, 34-35
Obligation, 226
Observation, 48, 261, 376-393, 483-484, 495
  of children, 391-393
  of dyadic situation, 384-387
  of face-to-face groups, 376-384
  in a field setting, 387-391
  in therapy situation, 385-387
Observation systems, category sets for, 372-375
  context of, 378, 380-381, 384
  dimensions of, 374-375
  orientation of the observer in, 377-383, 387, 389-390
  pre *vs.* post categorization, 399
  problems of, 396-400
  purpose of, 376-382, 384-393
  rating scales for, 372-375
  reliability of, 378, 380-382, 384-388, 391-393, 396-397
  sampling of behavior in, 398-399
  scores for, 378, 380-382
  typical systems, 375-393
    Bales', 376-378
    Carter's, 382-384
    Chapple's, 384-385
    conference research problem solving, 378-380
    self-oriented need, 381-384
    Steinzor's, 380-381
  unit of observation, 375, 378-385, 387, 391
  validity of, 397-398
Observational techniques, 370-403
  equipment for, 393-396
  history of, 370-371
  occasions for, 371-372
  optimal conditions for, 400-402
  and psychophysics, 400-402
Observer, effect of, 399-400
  tasks of, 372-373
Oedipus complex, 153-154, 166
Open-ended question, 341
Operational analogy, 80-81
  extension of theory, 79
Operationism, 57-58, 91-92, 107-108, 188, 264
Operating characteristics, 289-303, 343, 350-352, 359-360
  definition of, 291
  of items, 343, 350-352, 359-360
Opinion polling, 18, 48, 341
Opinion statement, 341
Organismic variable, 263, 269
Organization, level of, 131-133, 136
  manifest, 103-104
  perceptual, 97-99, 104-107
  silent, 109-114
  See also: cognitive organization
Out-of-the-field behavior, 206
Overlapping situation, 198, 206-207
Paired comparison scaling, 345-347
Parameter, 266, 332-333
Parent figure, 243-244
Participation, 209
Past experience, 106, 108-109, 208-209
Path, 189, 196-197, 199
Perception, 36, 97-99, 104-107, 116, 151-152, 210-211, 228-231, 241, 253-255, 544-545
  and motivation, 116
  of persons, 106-107, 210-211, 228-231, 241, 253-255, 544-545
  and psychoanalysis, 151-152
  social, 36, 104-107, 210-211
Perceptual representation, theories of, 99-103
Perceptual unit formation, 98-99, 103-107
Person, as cause, 106-107, 194
  in field theory, 190-191
Personality, 113, 148-155, 224, 239, 432-436
  and culture, 239
  effect of on cognition, 113
  and sociometric status, 432-436
  subdivisions of, 148-155
Personality measurement, and content analysis, 500-501
Personality scaling, 250-251
Phenomenal causality, 103-104
  field, theories of, 99-103
Phenomenology, 18, 20-21, 28-29, 92-93, 101-102, 122-123, 183-185
  definition of, 101-102
Phenotype, 187
Phobia, 160-161
Phrenology, 8
Physiognomic quality, 123
Play, 226-227
Pleasure principle, 149-150
Poisson distribution, 326
Position, in field theory, 196
  in a group, 218-219
  in role theory, 223-225
Positivism, 4, 7, 57-58, 91-93, 107-108
Postulate system, Hull's, 70
Posture, and role perception, 230-231
Potency, 197-198, 207
Power, 13-15
Power curve, 289-303, 314
  definition of, 291
Power field, 209-210
Pragmatism, 57-58
Preconscious, 155-156
Prediction, 266-267
Prejudice, 342-343, 425-426
  sociometric measures of, 425-426
Prestige, 47
Primal horde, 167
Primary group, 167
Primary process, 149-150
Primitive passive sympathy, 20-23

Principle, of stimulus-response contiguity, 74-75
  of utility, 10
Privacy, 102-103, 123
Probabilities, multiplication of, 329
Probing, nondirective, 466, 478
Problem solving, 71, 376-379
  by groups, 376-379
Process analysis, 78
Productive thinking, 109
Productivity, of groups, 233
Projection, 164, 237, 342
Projective questions, 460-462
Projective techniques, and content analysis, 499
  and sociometric status, 435-436
Propaganda, 28, 495-496, 501-502, 512
Propensity, 41
Propositional language, 125-127
Proximal determination, 95-97
Psychic actuality, 36-37
Psychoanalysis, and anthropology, 170-171
  applications of to other social sciences, 170-175
  origins of, 144
  and psychology, 143-145
  recent trends in, 168-170
  research methods of, 169-170
  and social psychology, 145, 165-168, 172-175
  and sociology, 171-172
  and stimulus-response theory, 58
Psychoanalytic theory, 143-175
  anxiety, 160-163
    and defenses, 162-163
    definition of, 160
    origins, 161
    in relation to ego, 162
    in relation to superego, 162
    types of, 160-161
  ego, 148, 150-155, 159, 162, 168-169
    defenses, 162-164
      other defenses, 164
      projection, 164
      reaction formation, 164
      regression, 164
      repression, 162-164
      sublimation, 164
    functions of, 150-151
    and memory, 151-152
    origins of, 150
    and perception, 151-152
    in relation to anxiety, 162
    in relation to id, 153
    in relation to instincts, 159
    in relation to superego, 154-155
    theory, recent trends, 168-169
    and thought and judgment, 152-153
  energy, 145-147, 153
    of ego, 153
  cathexis, 146-147, 156, 163
    and conscious-unconscious distinction, 156
    and repression, 163
  conscious, 148, 155-156
    preconscious and unconscious, 155-156
    and subdivisions of personality, 148
  defenses (*see* ego, defenses)
  id, 148-150, 153-154, 163
    contents, 148-149
    and ego, 153
    functions, 149-150
    origin, 148
    and primary process, 149-150
    and secondary process, 150
    and superego, 154
  instincts, 146, 156-160, 169
    characteristics of, 157
    death, 146, 156-160, 169
    definition of, 156-157
    and the ego, 159
    and the id, 148-149
    life, 146, 157-159
    number of, 157
    origin of, 156-157
    and the superego, 159-160
    theory, recent trends, 169
  libido, 146, 157, 169
  pleasure principle, 149-150
  preconscious, 155-156
  primary process, 149-150
  reality principle, 150-151, 153
  secondary process, 150-151, 153
  social psychology, 165-168
    culture, 167-168
    family structure and dynamics, 165-167
    group psychology, 167
    socialization of the individual, 165
    society, origins of, 167
  superego, 147, 153-155, 159-160, 161-162
    and anxiety, 161-162
    and ego, 154-155
    functions of, 154
    and id, 154
    and instincts, 159-160
    origins of, 153-154
  unconscious, 148, 151-152, 155-156
Psychological field, 185, 189-195
Psychology, social *vs.* physiological, 8
Psychometric function, 343
Psychopathology, and role theory, 233
Psychopathy, 245-248
Psychophysics, and observational techniques, 400-402
Psychosexual development, 158
Psychosis, 252
Psychotherapy, 385-386, 440
  and stimulus-response theory, 82-83
Punishment, 83, 205
Purpose, 94, 183
Purposive behavior, 62, 94, 538-539

Quantitative techniques, 289-333
Quasi-stationary equilibrium, 217
Question, direct *vs.* indirect, 460-462
  open *vs.* closed, 457-459
  phrasing of, 455-459, 476-477
Questionnaire, 338, 344, 482-483

Randomization, 269-271
Randomized blocks design, 278-279
Randomized groups design, 277-278
Range, 304, 323-324
Rank order scaling, 347
Rankings, multiple, 290, 319-320
Rapport, 25, 464-465, 474-475
Rating, 264
  errors in, 271-272
Rating scales, 272-275
  vs. sociometric measures, 406
Rationalism, 15-18
Rationalization, 14, 16-17
Reaction formation, 164
Readability, 496-497
Reality, level of, 192-193
Reality principle, 150-151, 153
Reality testing, 151
Recall of interrupted tasks, 200-202, 218
Reciprocal action, 223
Redintegration, 26
Reduction of determining tendencies, 27
Reductionism, genetic, 124-132
Reference scale, 208
Region, in field theory, 196-197
Regression, 157, 164, 203-204, 249-250
Regression effects, 279-281
Reinforcement, 11-13, 60, 114, 117-118, 120-121, 242
  and imitation, 23
Reinforcement theory, 63-71, 226
Relational analysis, 409
Relational determination, 98-99, 130
Reliability, 264, 338-340, 363, 371, 378, 380-382, 384-388, 391-393, 396-397, 420-422, 453-454, 514-515
  of content analysis, 514-515
  of interviews, 453-454
  of observation systems, 378, 380-382, 384-388, 391-393, 396-397
  of questions, 462
  rater, 363
  of scales, 338-340
  of sociometric measures, 420-422
Religious experience, 34
Repetition compulsion, 157
Replications, number of, 272-273
Representations, collective, 37-38
Repression, 31, 152, 161-164, 236-237
  and crowd behavior, 31
Research, types of, 259-260
Respondents, scoring of, 350
Response, cognitive definition of, 108-109
  fractional anticipatory, 62
  nature of, 73
Response-response theory, 77
Responsibility, 210-211
Resumption of interrupted tasks, 201-202
Right, 226
Rigidity, 136, 201, 247, 249-250
Rites de passage, 235
Role, 223-227, 232-233, 235-238, 248-255, 380, 462-464, 473
  accessibility of, 235-236
  acquisition of, 226-227
  "as if" dimension of, 236-238
  interaction with self, 248-255
  number of, 232-233
  relationships in interviewing, 462-464, 473
  vs. self, 253-255
Role conflict, 228
Role enactment, 232-238
  level of involvement in, 233-235
  organismic dimension of, 233-237
Role expectation, 225-229
Role perception, 229-232
Role-playing technique, 228
Role-role conflict, 252-253
Role taking, 232, 237-238, 243, 246-248
Role theory, 223-255
  origins of, 223-224
  vs. interaction theory, 223
Rorschach test, 249-250
Run test, 290, 321-323

Sampling, 265, 398-399, 515-516, 526-527
  in content analysis, 515-516
  in cross-cultural studies, 526-527
Satiation, 202-203
Saturation point, 425
Scale criteria, 363
Scale, interval, 364
Scaling, 264, 336-341, 344-365
  comparison of methods, 363-365
  judgment methods, 344-351
  latent class analysis, 360-361
  latent distance scale, 361-362
  latent structure analysis, 359-362
  maximum variance score, 353, 358
  method of summated ratings, 351-353
  principal components, 358
  rating methods, 362-363
  response methods, 344
  single common factor model, 352-353
  unfolding techniques, 358-359
Scalogram analysis, 353-358, 361-362
Schema, 110-111, 113, 116
Schizophrenia, 247
Science, classification of, 7-8
Secondary process, 150-151, 153
Secondary reinforcement, 62, 72, 115, 118-120

Selective distortion, 342
Self, 26, 29, 223-224, 238-255
  and cognitive tasks, 248
  empirical methodology for assessing, 244-245
  experimental study of, 245-248
  interaction with role, 248-255
  vs. role, 253-255
  subconscious, 26, 29
  See also: ego
Self-actualization, 115-116
Self-concept, 121-122
Self-consciousness, 236
Self-oriented need, 381
Self-role conflict, 251-252
Sentiment, 41, 45-46
  self-regarding, 46
  vs. attitude, 45-46
Set, 43-44, 110
  See also: aufgabe, determining tendency
Sex, and sociometric status, 430
Sex differences, in role expectation, 228-229
  in role perception, 230
Sex role, 226
Sexual instinct, 146, 157-158
Sibling rivalry, 166-167
Sign, 128-130
Sign gestalt, 104
Sign test, 290, 312-314
Signal, 128-130
Significance, tests of, 261-263, 268, 290-303, 331-332
  combining, 290, 328-331
  one-tailed *vs.* two-tailed, 262
Significant symbol, 22
Silent organization, 109-114
  See also: set
Similarity, association by, 37
Simple and sovereign theories, 9-29
Simplicist fallacy, 9
Simplicity, 125, 132
Sleeper effect, 84
Small group, co-acting, 47
  early studies of, 46-48
  face-to-face, 47
Small group theory, 232-233, 376-384
Social act, 535, 545-548
Social class, 86, 228
Social climate, 48
Social contract, 13, 15-16
Social disorganization, 252
Social distance scale, 353
Social facilitation, 47
Social increment, 47
Social norm, 195, 253
Social perception, 36, 104-107, 210-211
Social psychology, and abnormal psychology, 25-26
  Aristotelian, 6
  Comtean, 7-9
  definition of, 4-5
  early experimentation in, 46-48
  founders of, 3, 7-9
  history of, 3-51
  objective methodology in, 46-48
  origins of, 3-4, 6-9
  Platonic, 6
  and psychoanalysis, 145, 165-168, 172-175
  secondary sources for history of, 5-6
  and sociology, 4-5, 50-51
  textbooks in, 48-51
  units of analysis of, 40-46
Social sensitivity, 409, 427
  See also: empathy
Social stimulus value, 392-393
Social subvaluent, 47
Socialization, 15, 83-84, 153-154, 165
  and egoism, 15
Socially induced change, 209-210, 217
Society, 31-32, 167, 224, 552-558
  conditions for, 31-32
  insect, 535-537
  origins of, 167
Sociology, 4-5, 7-9, 38, 171-172
  Comte's invention of, 7-9
  of knowledge, 17
  and social psychology, 4-5
Sociogram, 410-413
  appraisal of, 412-413
  target, 412
Sociometric choice, intensity of, 414
  patterns of, 410-411
Sociometric data, analysis of, 410-420
  factor analysis of, 419
  fractionation of groups of, 419-420
  graphic representation of, 410-413
  indices from, 413-414
  machine computation of, 419
  matrix approaches to, 416-419
  profiles for, 415
  scores from, 413-414
  simple quantitative methods for, 413-415
  statistical methods for, 415-416
  tabular methods for, 413
Sociometric measures,
  applications of, 436-441
    community applications, 440-441
    educational applications, 436-438
    industrial applications, 438
    military applications, 439-440
    psychotherapeutic applications, 440
  data appropriateness of, 423
  definition of, 407-408
  empirical concepts definable by, 424-429
  internal consistency of, 421
  problems of, 441-444
  qualities of, 405-406
  and related devices, 409-410
  relation to independent measures of, 423

  - reliability of, 420-422
    - interpretive reliability of, 420-421
    - and multiple criteria, 421-422
    - repeat reliability of, 421
    - test reliability of, 421
  - research utility of, 423-436
  - validity of, 422-424
    - face validity of, 422-423
  - variations of, 408-409
- Sociometric self-rating, 409
- Sociometric survey, multirelational, 409
- Sociometric variables, measures related to, 429-436
  - cultural variables, 430-431
  - demographic variables, 429-430
  - personality, 432-436
- Sociometry, 48, 405-444
  - origins of, 407
  - use of the term, 405
- Socius, 242
- Soviet psychology, 34, 47-48
- Spontaneous recovery, 61
- Square root transformation, 326
- Stability, 136, 250
- Standard deviate, 329
- Standard deviation, of a mean, 323
- Statistical techniques, 289-333
  - quick methods in, 290, 323-324
- Status, 224, 427-430
  - sociometric measures of, 427-428
  - and sociometric variables, 429-430
- Stereotype, 17
- Stimulus, equivalence, 452-453
  - generalization, 61-62, 75, 130-132
  - nature of, 73, 97-99
  - *vs.* need, 156-157
- Stimulus-intervening variable law, 67
- Stimulus-response theory, 57-87
  - categorization of behavior sequence, 59-61
  - *vs.* cognitive theory, 93, 96-97, 113-115, 118-120
  - contiguity theory, 63, 73-77
  - dual theories, 64-65, 71-73
  - and language, 129-130
  - origins of, 57-58
  - and other behavior sciences, 85-87
  - and reductionism, 124-125
  - reinforcement theory, 63-64, 65-71
  - and social phenomena, 69-70, 81-85, 119-120
  - systems of, 59, 63-77
- Stimulus-stimulus theory, 70-71
- Stimulus variable, 66
- Stochastically smaller scores, 317-319
- Structural analysis, 78
- Subconscious self, 26, 29
- Subjects, number of, 266-267, 277, 292-294, 302-303, 330-331
  - randomization of, 269-270
  - selection of, 265-266, 273-274
- Sublimation, 164
- Submission, 24, 391-392
- Substitution, 202
- Successive intervals, method of, 347-349
- Suggestibility, 25-27, 29, 237
  - conditions of, 27, 29
- Suggestion, 18-19, 24-30
  - definition of, 24, 26-27
  - explanations of, 24-29
- Sum of squares, 267-268
- Superego, 147, 153-155, 159-162
- Survey research, 259
- Symbol, 128-130
  - significant, 22
  - universal, 34
- Symbolic function, 121, 125-130
- Sympathy, 18-23
  - empirical studies of, 21
  - forms of, 19-21
  - primitive passive, 20-23
  - proponents of, 19-20
- Symptoms, 163
- Syncretism, 136
- System, personality *vs.* social, 40
  - theoretical, 59-77
  - theoretical-empirical, 59, 77-87

- T statistic, 315-317
- t test, 272-273, 303, 323-324
- Taboo, 556-557
- Task orientation, 122, 202
- Technique research, 260
- Tension binding, 239, 245-246
- Tension reduction, 11-13, 114-115, 240-241
- Tension system, 112, 199-205
- Testing, 264
- Textbooks in social psychology, 48-51
- Theory, problem of generalizing, 77-81
  - simple and sovereign, 9-29
- Thinking, 125-130, 152-153
- Thurstone type scale, 344-351
- Time binding, 239
- Topological diagram, 67, 198-199
- Topology, 195-199
- Torres Straits Expedition, 35
- Total situation, 185-186, 189-195
- Trace line, 343, 359
  - system, 111-113
- Trait, 238
- Transfer of training, 75, 109, 131-132
- Transformation of motives, 15, 70-71
- Transformations, 290, 324-328
- Transitoriness of instinct, 42
- Transposition, 130-132
- Trauma, 160-161
- Triadic hypothesis, 236-237

- U statistic, 290, 317-319
- Unconscious, 109-111, 148, 151-152, 155-156

Unidimensionality, of a scale, 339-340, 363
Unitary explanation, fallacy of, 9
Utility, principle of, 10

V statistic, 317-319
VTE, 245
Valence, 198, 200, 204, 206, 208
Validity, of content analysis, 504-505
  of interviews, 454, 479-482
  of observation systems, 397-398
  of questions, 458, 460-462
  of scales, 338, 340-341
  of sociometric measures, 422-424
Value, 45, 113-114, 431-432, 510
Variable, 263-264
  independent *vs.* dependent, 261
  quantification of, 264
Variance estimate, 329-330
  stabilizing transformations, 325-326
Variation between groups, 265, 268-272
Veridicality of role perception, 229-230
Vocalization, 126
Voodoo death, 235
Vulnerability of a theory, 65

W-A-Y technique, 244-245
Wagging dance, 537-538
Wald-Wolfowitz run test, 290, 321-323
Weaning, age of, 239, 524-525
Whitney's U statistic, 290, 317-319
Wilcoxon's signed rank test, 290, 314-315
Wish, 149-150
Würzburg school, 44

Form No. 3.

PSY, RES.L-1

Bureau of Educational & Psychological Research Library.

The book is to be returned within the date stamped last.

12.4.60		
23.3.61		
- 2 JUN 1962		
22.2.74		

WBGP-59/60-5119C-5M

301.1502
LIN

Form No. 4

BOOK CARD

Coll. No. 301.1502

Author Lindzey Gardner Accn. No 1268

Title Hand book of Social Psychology

Date.	Issued to	Returned on
12.4.60	260	23.6.60
23.3.61	K.8.C	4 MAR 1961
2 JUN 1962	RCJ	
22.2.74	B.Sg.	

301.1502
LIN

Form No. 3.

PSY, RES.L-1

Bureau of Educational & Psychological Research Library.

The book is to be returned within the date stamped last.

WBGP-59/60-5119C-5M

Form No. 4
BOOK CARD

Coll. No.............. Accn. No......

Author..

Title..

Date.	Issued to	Returned on
....	....	....